UNIVERSITE JOSEPH FOURIER – GRENOBLE 1

ECOLE DOCTORALE DE PHYSIQUE

THESE

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE JOSEPH FOURIER
Discipline: Physique

présentée par

Paul SCHANDA

Développement et applications de méthodes RMN rapides pour l'étude de la structure et de la dynamique des protéines

Development and application of fast NMR methods for the study of protein structure and dynamics

Directeur de thèse: Dr. Bernhard BRUTSCHER

Thèse soutenue publiquement le 9 octobre 2007 devant le jury:

Dr. Bruno KIEFFER	Rapporteur
Dr. Martin TOLLINGER	Rapporteur
Dr. Vincent FORGE	Examinateur
Dr. Lucio FRYDMAN	Examinateur
Dr. Judith PETERS	Examinateur
Dr. Nico VAN NULAND	Examinateur
Dr. Bernhard BRUTSCHER	Directeur de thèse

Gewidmet meinem Bruder Gregor

Remerciements

Ces trois années de thèse ont été pour moi un très grand plaisir, et je peux dire que j'avais envie tous les jours de retrouver mon travail et surtout les gens avec qui j'ai travaillé. Je pense que c'est une chance extraordinaire, et je tiens à remercier toutes les personnes qui ont fait de cette thèse une véritable source d'épanouissement personnel et professionnel et qui créent cette ambiance à la fois motivée, scientifiquement excellente et aussi détendue et agréable.

Tout d'abord, je voudrais remercier Bernhard Brutscher, mon guide scientifique tout au long de ces années. J'ai apprécié non seulement ses compétences scientifiques et son efficacité impressionnantes (il fallait "s'accrocher" au début pour pouvoir suivre...) mais aussi son humour et l'ambiance détendue dans laquelle il m'a fait découvrir la RMN. Danke, Chef-Spektroskopiker !

Je tiens à remercier Jean-Pierre Simorre. En tant que chef du LRMN il m'a permis de profiter d'une très bonne infrastructure pour la science, et d'un environnement humain agréable. Ce RMNiste à grande 'largeur spectrale' scientifique m'a appris bien des choses sur la RMN, toujours avec un regard critique sur mon travail et toujours en plaisantant ("c'est tout con, la sofast..."). J'ai eu aussi le plaisir de partager pas mal de belles ballades en montagne avec lui.

Je tiens également à remercier mes collègues de bureau, Julien Boudet et Beate Bersch, avec qui j'ai pu partager les hauts et bas d'un spectroscopiste, et qui m'ont beaucoup aidé scientifiquement.

Merci aussi à Ewen Lescop, avec qui j'ai eu le plaisir de travailler, de façon très efficace et détendue, et qui m'a appris beaucoup de choses sur la RMN.
Pour l'excellente collaboration pendant son stage, je voudrais remercier Mirjam Falge, qui a - avec son esprit vif et critique - contribué beaucoup au succès du projet RDC-HNHN.

Pendant quelques semaines, j'ai travaillé avec Isabel Ayala et Cécile Giustini au PSB pour produire une protéine. Je garde un très bon souvenir de cette première expérience en biochimie, qui m'a montré aussi les étapes parfois difficiles avant d'avoir une protéine dans le spectromètre (on n'y pense pas toujours en tant que spectroscopiste...).

Pour leur disponibilité à discuter de nombreuses questions scientifiques, je remercie tous les gens du LRMN, en particulier Dominique Marion (avec ses comparaisons qui rendent la RMN plus compréhensible), Beate Bersch, Catherine Bougault (qui est toujours prête à partager ses très grandes connaissances dans divers domaines scientifiques), Martin Blackledge, Pierre Gans (qui m'a aussi montré des belles ballades en vélo), Klaartje Houben, Jérôme Boisbouvier, Adrien Favier et Ewen Lescop.

Merci à Adrien aussi pour les nombreuses fois où il m'a aidé avec mes problèmes informatiques et son excellent travail de maintenance des spectromètres, qui facilite énormément la vie pratique des spectroscopistes au LRMN.

Je tiens aussi à remercier notre groupe de thésards et post-docs, Julien Boudet, Rémy Sounier, Guillaume Bouvignies, Monika Nalezkova, Thomas Kern, Nico Duraffourg, Klaartje Houben, Fito Rasia, Pau Bernado, Matthew Devany, Malene Ringkjobing-Jensen, Hélène Van Melckebeke, Florence Mannon, Cedric Laguri et Ewen Lescop. On était vraiment une très bonne équipe. Merci pour toutes les soirées que l'on a passées ensemble et pour les interactions scientifiques.
Merci aussi à Phineus Markwick, qui a essayé pendant des soirées entières de m'initier à la Density Functional Theory (et plus tard dans la soirée aussi aux secrets du whiskey écossais).

Comme la montagne est une grande source d'inspiration et de 'relaxation', je remercie les personnes avec qui j'ai découvert les montagnes grenobloises et qui étaient très importantes pour moi pendant ces années: Nico Duraffourg (l'homme qui trouve toujours les chemins les plus raids), Adrien Favier, Klaartje Houben, Thomas Kern et mes colocataires Julien Claudon, Karol Marty, Tobias Robert et François Chamaraux.
Un grand merci aussi à Célia Bonnet pour notre temps au 124 bis Cours Berriat, Nico et Stanislas Georges, Elisabeth Weber, Olivier Seneque, Raimond Ravelli, Vicky Markwick ...

Je voudrais remercier spécialement Hélène Van Melckebeke, qui était très importante pendant ces années et qui est responsable du fait que je me suis retrouvé

à Grenoble.

I also had the chance to spend three months in Israel to work at the Weizmann institute (merci à Bernhard et Jean-Pierre de m'avoir donné cette liberté). Both scientifically and personally this time was very enriching for me, and I would like to thank especially Lucio Frydman and Maayan Gal for their hospitality and a very stimulating and interesting collaboration. Thanks to Maayan also for some unforgettable hikes in some of Israel's nicest places.

I would also like to thank Bruno Kieffer and Martin Tollinger, who have accepted to review this manuscript, as well as the examiners Vincent Forge, Lucio Frydman, Judith Peters and Nico Van Nuland. Pour la relecture de ce document je remercie Bernhard, Ewen, Catherine et Hélène.

Contents

List of Figures

List of acronyms and symbols

afb	aliphatic flip back
BEST	Band Selective Short Transient
BURP	Band-selective Uniform Response Pure-phase: family of selective rf pulses: containing excitation (EBURP) and refocussing (REBURP) pulses
CW	Continuous Wave
DNP	Dynamic Nuclear Polarization
EBURP	see BURP
EPI	Echo Planar Imaging
EX1	Limit of amide hydrogen exchange that gives direct access to opening rates
EX2	Limit of amide hydrogen exchange that gives access to equilibrium constants of opening/closing equilibria relevant for hydrogen exchange
FID	Free Induction Decay
FT	Fourier Transform
HMQC	Heteronuclear Multiple Quantum Coherence
HNCO	Experiment that correlates the resonance frequencies of amide proton, nitrogen and the neighboring carbonyl atom
HSQC	Heteronuclear Single Quantum Coherence
INEPT	Insensitive Nuclei Enhanced by Polarization Transfer
IPAP	In Phase-Anti Phase
ISNOB	see SNOB
Jcomp	J-mismatch compensation: pulse sequence element for coherence transfer via a large band of J-couplings
J	Scalar coupling between two spins
LP	Linear Prediction
nD	n-dimensional, used in the context of multidimensional NMR
NMR	Nuclear Magnetic Resonance
nOe	nuclear Overhauser effect

PC	Polychromatic Pulses (Kupce and Freeman (1994))
PDB	Protein Data Bank: database for structural coordinates of biomolecules
ppm	parts per million: unit used to describe the chemical shift
RDC	Residual Dipolar Coupling, induced by a partial alignment of molecules in solution
REBURP	see BURP
rf	radio frequency
se	sensitivity enhanced (used in terms of preservation of equivalent pathways, as introduced by Cavanagh and Rance (1990)
SiR-FP18	Flavodoxin domain of the sulfite reductase (SiR), molecular weight 18kD
SNOB	Selective excitatioN fOr Biochemical applications: family of selective excitation, inversion and refocussing pulses
SOFAST	band-Selective Optimized Flip Angle Short Transient
SS-HMQC2	Semi-Selective double HMQC experiment
TROSY	Transverse Relaxation Optimized SpectroscopY: technique that exploits cross-correlated relaxation between two relaxation mechanisms to enhance sensitivity and/or resolution
wfb	water flip back

Résumé
Développement et applications de méthodes RMN rapides pour l'étude de la structure et de la dynamique des protéines

La RMN multidimensionnelle (nD) est la méthode de choix pour l'étude structurale et dynamique des protéines en solution avec une résolution atomique. Une limitation de la RMN-nD est la longue durée de l'acquisition de données nD: le temps d'acquisition du jeu de données nécessaire pour une étude structurale est souvent de l'ordre de plusieurs semaines. De plus les processus cinétiques, qui se passent à l'échelle de la seconde, ne sont pas accessibles aux études en temps réel par RMN-nD en utilisant les méthodes standards. Ce travail présente des développements méthodologiques qui visent à accélérer la RMN-nD en optimisant la relaxation longitudinale des protons amides. Les méthodes proposées permettent d'acquérir des spectres de corrélation 2D ^{1}H-^{15}N (3D ^{1}H-^{15}N-^{13}C) en quelques secondes (quelques minutes). En plus, en combinaison avec des méthodes existantes (encodage spatial, encodage Hadamard), le temps d'acquisition pour des spectres 2D peut être réduit à une seconde. Des applications à l'étude de phénomènes cinétiques des protéines sont presentées. Cette thèse présente également une nouvelle expérience qui permet d'évaluer rapidement la qualité d'un échantillon de protéine, et une méthode pour mesurer des couplages résiduels dipolaires entre protons amides avec une meilleure sensibilité que les méthodes existantes.

Mots Clés: RMN multidimensionnelle, relaxation longitudinale, Effect Overhauser Nucléaire, Couplage Résiduel Dipolaire, repliement de protéines, échange H/D des protons amides, impulsions sélectives

Laboratoire de thèse: Institut de Biologie Structurale Jean-Pierre Ebel, UMR 5075, C.E.A.; C.N.R.S.; U.J.F. Laboratoire de Résonance Magnétique Nucléaire. 41 rue Jules Horowitz, 38027 Grenoble Cedex, France.

Abstract
Development and application of fast NMR methods for the study of protein structure and dynamics

Multidimensional (nD) NMR is the method of choice for atom-resolved studies of protein structure and dynamics in solution. Among its current limitations are the long acquisition times required, translating to experimental times of several days or weeks for the set of experiments required for structural studies of proteins. Furthermore, real-time studies of kinetic processes occurring on a seconds time scale are inaccessible to standard nD NMR. This thesis is concerned with the development of fast nD NMR techniques based on longitudinal relaxation optimization. It is shown that 2D ^{1}H-^{15}N (3D ^{1}H-^{15}N-^{13}C) correlation spectra can be obtained in only a few seconds (few minutes) of acquisition time for samples at millimolar concentration. In addition, the longitudinal relaxation optimized methods, when combined with alternative data sampling such as spatial or Hadamard encoding, can yield site-resolved 2D ^{1}H-^{15}N correlation spectra in acquisition times down to one second. Applications of fast 2D methods to the study of protein folding and unfolding are shown. This thesis also presents a longitudinal relaxation optimized method for the sensitivity-enhanced measurement of residual dipolar couplings between amide protons, as well as a fast and simple experiment for characterizing protein samples, which can be very useful in the context of screening of sample conditions.

Key Words: Multidimensional NMR, longitudinal relaxation, Nuclear Overhauser Effect, Residual Dipolar Coupling, protein folding, amide proton exchange, molecular kinetics, selective pulses

Résumé: Méthodes rapides pour l'étude des protéines par RMN

Introduction

Les principaux acteurs des processus biologiques sont les protéines et les acides nucléiques. Pour comprendre leur fonctionnement au niveau moléculaire, il est souvent indispensable de connaître leur structure tri-dimensionnelle, ce qui est le sujet de recherche de la Biologie Structurale.

Parmi les méthodes qui se prêtent à une étude structurale de biomolécules, la spectroscopie par Résonance Magnétique Nucléaire (RMN) joue un rôle crucial, car elle est pour le moment la seule technique qui permet de résoudre la structure de biomolécules en solution avec un résolution atomique. Au cours des 20 dernières années, la RMN s'est donc établie comme la principale méthode pour l'étude de la structure, la dynamique et des interactions biomoléculaires en solution.

Un pas crucial dans le développement de la RMN, qui a permis son grand succès, a été l'introduction de la RMN multidimensionnelle (nD: n égal à 2, 3,...). La RMN nD a permis d'augmenter la résolution des spectres RMN par une séparation des signaux dans un espace multidimensionnel, en corrélant les fréquences de résonance de plusieurs noyaux atomiques (par exemple la fréquence de résonance du noyau de l'azote amide et celle du proton directement lié). Cette augmentation de résolution a permis de passer de l'étude de petites molécules vers celle de plus grandes molécules, comme les protéines.

Un problème associé à la RMN nD est lié au temps nécessaire pour collecter un spectre nD. Pour un spectre 2D, ce qui est le minimum pour résoudre les signaux qui correspondent à des atomes individuels dans une protéine, le temps

expérimental nécessaire est de l'ordre de quelques minutes. Le temps expérimental nécessaire pour l'acquisition de spectres en 3D est de l'ordre de quelques heures, et la durée expérimentale augmente de façon exponentielle avec la dimensionalité.

Ces longues durées expérimentales ont des conséquences pratiques pour les études RMN: Pour le jeu de spectres nécessaires à une étude structurale d'une protéine, le temps d'acquisition est souvent de l'ordre de plusieures semaines, ce qui limite l'utilité de la RMN dans le contexte d'études à haut débit (comme par exemple dans des études de génomique structurale). En plus, et plus fondamentalement, les méthodes standards de RMN multidimensionnelle, avec des temps d'acquisition de l'ordre de plusieures minutes au minimum, sont incapables de suivre en *temps réel* des processus qui se passent à l'échelle de la seconde. Par exemple, les méthodes standards ne permettent pas de suivre les changements spectraux au cours du repliement de protéines, un processus qui se passe à l'échelle de la milliseconde jusqu'à la minute.

En vue de ces limitations, les dernières dix à quinze années ont vu un effort considérable afin de réduire le temps expérimental pour la RMN nD. Le travail présenté ici s'inscrit dans ce contexte.

Résultats et Conclusion

Pour l'acquisition de spectres multidimensionnels RMN, d'après le schéma classique, il est nécessaire de répéter un élément de base (un 'scan') de multiples fois. Entre deux répétitions, un délai d'attente est nécessaire pour rééquilibrer le système (par un processus que l'on appelle *relaxation longitudinale*). L'accélération de l'acquisition de spectres nD est donc possible à deux niveaux: soit le nombre de répétitions nécessaires est diminué, soit le délai d'attente est raccourci.

Differentes méthodes qui visent à diminuer le nombre de répétitions ont été proposées dans la littérature (voir chapitre 2.2). Dans ce travail, nous avons suivi l'autre stratégie: la diminution du délai d'attente entre deux scans. Pour celà, la relaxation longitudinale des spins nucléaires étudiés doit être optimisé. Nous avons implementé des expériences qui manipulent de façon sélective seulement un groupe de protons de la protéine, ce qui permet une relaxation plus efficace et donc une meilleure sensibilité. Le chapitre 2.3 explique en détail les principes de ces expériences.

Une première expérience basée sur une optimisation de la relaxation longitu-

dinale, que nous avons appelée SOFAST-HMQC (ce qui signifie band-Selective Optimized Flip Angle Short Transient Heteronuclear Multiple Quantum Coherence) permet d'acquérir des spectres 2D en seulement quelques secondes pour des protéines à une concentration de l'ordre d'un millimolaire. Ce type de spectre 2D, qui corrèle la fréquence du proton amide avec celle de l'azote amide ou la fréquence des protons méthyles avec la fréquence du carbone lié, est à la base de la majorité des études de structure, de dynamique et d'interaction. La diminution importante du temps expérimental permet donc d'accélérer considérablement une partie des études par RMN. En outre, cette technique permet de suivre des changements spectraux pendant des processus cinétiques, comme le repliement de protéines, avec une résolution temporelle de quelques secondes.

L'expérience SOFAST HMQC est optimisée en terme de relaxation longitudinale, ce qui permet de répéter l'expérience sans délai d'attente. En combinaison avec des méthodes existantes qui visent à minimiser le nombre de répétitions nécessaires, le temps d'acquisition peut encore être reduit. Une expérience SOFAST HMQC avec un encodage spatial des fréquences, que nous avons mis au point en collaboration avec Maayan Gal et Lucio Frydman, de l'institut Weizmann, permet de réduire le temps d'acquisition pour un spectre 2D à moins d'une seconde dans des cas favorables (voir chapitre 7.2).

Les concepts utilisés pour la spectroscopie à deux dimensions ont été appliqués à des expériences à plus haute dimensionalité: Le chapitre 5 présente des approches qui permettent de réduire le temps expérimental nécessaire pour l'acquisition de spectres 3D à quelques minutes (type HNCO, HNCA etc.). Ce type d'expérience est utilisé pour l'attribution des signaux, et la réduction du temps expérimental que les nouvelles expériences offrent se traduit par une accélération de cette étape clé de toute étude structurale de protéine.

De plus, les expériences qui sont optimisées en terme de relaxation longitudinale permettent non seulement de réduire le temps expérimental en gardant une bonne sensibilité, mais aussi d'augmenter la sensibilité. Nous avons mis au point une expérience dédiée à la mesure de couplages résiduel dipolaires entre protons amides. Ce type d'expérience donne des informations sur l'angle et la distance entre les protons, et contient donc des informations structurales utiles. La plus grande sensibilité offerte par la nouvelle expérience proposée ici, comparé aux méthodes existantes, permet d'obtenir des informations plus précises dans le contexte de la détermination de structure.

Les propriétés de relaxation longitudinale des protons amide dans les protéines

peuvent aussi être exploitées pour obtenir de l'information structurale semi quantitative. Le chapitre 8 détaille une méthode simple et rapide qui, en utilisant des critères de relaxation, permet de caractériser un échantillon, et d'évaluer rapidement quelles sont les parties structurées ou flexibles dans une protéine.

Le chapitre 9 montre deux applications des nouvelles méthodes à l'étude de repliement /dépliement de protéines en temps réel. Les expériences rapides permettent de suivre, résidu par résidu, des changements pendant un processus cinétiques.

Le repliement de la protéine α-lactalbumine d'un état partiellement déplié vers létat natif a été suivi en temps réel après une initiation rapide de la réaction. Il était possible de caractériser simultanément les changements structuraux pour la plupart des résidus de la protéine. En analysant les changements d'intensité nous avons trouvé une même vitesse de repliement pour tous les résidus en suivant l'apparition de pics. En outre, pour la première fois, nous avons observé la diminution de l'intensité des pics qui correspondent à l'état déplié, et ce changement se passe à une vitesse similaire. Ces observations sont compatibles avec un chemin de repliement comportant un seul état de transition. Ces données nous permettent d'exclure la présence d'un état intermédiaire dans le repliement.

Les méthodes rapides nous ont permis aussi d'étudier l'échange H/D des protons amides avec le solvant dans la protéine ubiquitine dans des conditions où la vitesse d'échange donne de l'information sur la vitesse de dépliement. Nos données montrent une hétérogénéité des vitesses de dépliement, avec la partie C-terminale de la protéine qui se déplie plus rapidement que la partie N-terminale.

Pour conclure, de nouvelles méthodes RMN qui permettent d'une part d'accélérer l'acquisition des données nécessaires pour des études structurales, et d'autre part de suivre des cinétiques moléculaires en temps réel, ont été présentés. Ce travail constitue un avancement qui permet de rendre la RMN plus efficace et utile au service de la biologie structurale.

1 Nuclear magnetic resonance spectroscopy: a powerful tool for structural biology

Molecular interactions are the basis of biological processes

In his book "What is life?"[1], published in 1944, the physicist Erwin Schrödinger raised a question, which should keep scientists busy for decades since then:

"How can the events in space and time which take place in a living organism be accounted for by physics and chemistry ?" At that time, biology and physics were still relatively far from each other, but Schrödinger predicted:

"The obvious inability of present-day physics and chemistry to account for such events is no reason at all for doubting that they can be accounted for by those sciences."

In fact, the contact of Biology and Physics has triggered the development of scientific branches that have perhaps been the most active fields of science in the last 50 years. These bio-physical/bio-chemical disciplines aim at understanding the molecular (i.e. microscopic) basis of the macroscopic aspects of living organisms. For such a molecular description, one has to identify the actors inside the cell (mainly proteins and nucleic acids), and study how they interact with each other.

Proteins and nucleic acids are linear chains composed of a well-defined sequence of basic building blocks (amino acids, nucleotides), and this sequence determines the function of the molecule. In fact, a large number of weak interactions between chemical moieties along the chain dictate the three-dimensional structure (i.e. the shape) that the molecule adopts in solution. A particular, well-defined structure is generally a prerequisite for the molecule to function properly, e.g. to

[1]Schrödinger, E. (1944), *What is life ? The Physical Aspect of the Living Cell.*, Cambridge University Press

interact with other molecules.

A detailed understanding of the molecular function and interactions therefore requires knowledge of the three-dimensional structure of the involved molecules, and understanding the link between the atomic structure of a biomolecule and its function is the research subject of Structural Biology.

Because of the small size of these molecules, on the order of 10^{-9} metres, they escape direct observation, e.g. by optical microscopy. To date only two methods are capable of giving atomic details of biomolecular structures: Nuclear Magnetic Resonance spectroscopy and diffraction of X-ray (or neutron) irradiation by crystals of (bio-)molecules.

Nuclear Magnetic Resonance spectroscopy applied to biomolecules

NMR spectroscopy is based on a quantum mechanical property of the nuclei named nuclear spin. Many of the atoms found in biological macromolecules (in general several hundreds of atoms even in small proteins or nucleic acids) have a non-zero nuclear spin, and in NMR experiments each of these atomic nuclei acts as a *local spy* inside the molecule, reporting on the local structure and dynamics that it senses.

Collecting information from a large number of local spies allows one to reconstruct a global, coherent picture of the structure and dynamics of the molecule, or of the assembly of interacting molecules. The model that one reconstructs has to be in agreement with all the information that the individual atomic nuclei provide, and the more information from individual sites is available, the more precise can such a model be. Retrieving the signal of as many *individual* nuclei as possible is therefore crucial.

NMR's capabilities to give very detailed insight into the structure and dynamics of biomolecules may be exposed with two examples, shown in figure 1.1.

Panel (a) represents the three-dimensional structure of a complex of a protein with one of its physiological binding partners. The protein molecule U2AF^{65}-RRM3, and the peptide SF1 are important actors in the splicing of pre-mRNA, a key step in eukaryotic gene expression. The structures were obtained taking into account hundreds of distances between individual nuclei, and the position of each atom in the complex is defined with a precision on the order of one Ångström (10^{-10} m).

The information that one can obtain is not limited to the three-dimensional structure. Figure 1.1(b) gives an example of how NMR can yield insight into dynamics of a protein, i.e. transient fluctuations of its structure. Shown is the protein

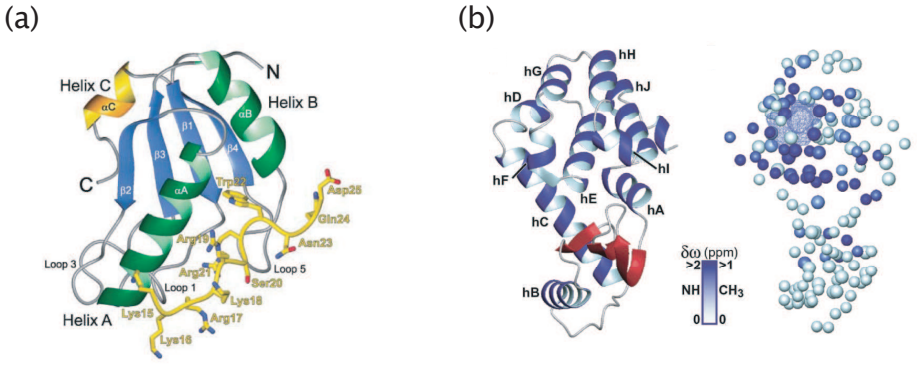


Figure 1.1 Two representative examples underlining NMR's capabilities for the study of biomolecular structure and dynamics. (a) NMR-derived structure of the protein U2AF^{65}-RRM3 in complex with the peptide SF1 (shown in yellow in the front). For simplicity, only the polypeptide backbone is shown in a so-called ribbon representation (Selenko et al. (2003)). (b) Identification of slow dynamics in the protein lysozyme T4 L99A variant (Mulder et al. (2001)). A cartoon of the three-dimensional structure is shown in the left panel. The binding site is located in the upper region of the molecule, as indicated by a blue grid in the right panel. Spheres in the right panel correspond to amide nitrogen atoms and methyl carbon atoms for which nuclear magnetic relaxation data are available. The color coding in the right panel displays the difference of the chemical shift (see following chapter) between the ground state and the excited state: purple spheres indicate that the chemical shift of the corresponding nucleus in the excited state deviates significantly from the one in the ground state. As the chemical shift is sensitive to local structure, these data indicate that the structure of the excited state differs from the ground state mainly around the binding site, and points to an excited state that is structurally different from the ground state especially around the binding site.

lysozyme T4 L99A, which is known to bind benzenes and substitutes thereof in a binding site located at the top of the molecule (indicated in the right panel). Interestingly, when considering the structure as determined by NMR or X-ray crystallography, the entry of a benzene molecule into this site is sterically impossible. Therefore, the knowledge of the three-dimensional structure alone cannot explain why the protein is able to fulfil its function. A phenomenon called nuclear magnetic relaxation, which is again sensed by individual atomic nuclei, gives more insight: Such relaxation experiments reveal that there is significant dynamics around the binding site, and point to an excited state of the protein molecule, in which the entry of the substrate molecule is possible. In this case, an understanding of the protein function is only possible with the knowledge of both the protein structure and dynamics, both of which are accessible to NMR methods.

To conclude, NMR proves a very powerful tool towards an understanding of biological processes on an atom-resolved basis.

Outlook

Obtaining detailed information about biomolecules by NMR requires the collection of a large amount of information from individual nuclei. The work presented in this thesis focuses on the development of experimental schemes that make certain aspects of this data collection procedure faster and thus more efficient, and applies the newly developed methods to the study of structure and dynamics of proteins.

Chapter 2 briefly reviews some of the relevant theory and introduces NMR techniques that allow obtaining signals from individual nuclei in proteins. The problems of current techniques in terms of the experimental time needed for NMR data collection are identified and strategies to solve them are discussed.

Methods for implementation of *longitudinal relaxation optimized* protein NMR spectroscopy are reviewed in chapter 3. Our own methodological developments done during my thesis are then presented in chapters 4 to 9.

Chapter 4 introduces a new experiment that allows to obtain 2D correlation spectra, displaying one peak per amino acid in the protein, in experimental times of only a few seconds. This concept is then extended in chapter 5 to triple-resonance experiments, that are needed for resonance assignment of backbone atoms in proteins.

A new and very sensitive method for the collection of structural information in proteins (amide ^{1}H-^{1}H residual dipolar couplings) is then proposed in chapter 6.

Chapter 7 shows the combined implementation of longitudinal relaxation optimization and alternative data sampling schemes, and proposes two methods that allow obtaining information simultaneously for a large number of sites in a protein molecule in an acquisition time of approximately one second.

In addition, a new technique for the rapid characterization of the structural compactness of protein samples is introduced in chapter 8.

Finally, the developed fast NMR techniques were applied to the real-time study of protein folding and unfolding reaction kinetics, which is demonstrated in chapter 9.

Large parts of the results of chapters 4 to 9 have been published in scientific

journals. In the present thesis I show some unpublished details, but I skip some of the published parts that I do not estimate necessary or not up-to-date any more. The references of the publications can be found on page 221. In addition, the names of the persons that were also involved in the different projects can be found in the list of publications. Especially, section 5.2 is to a large portion the work of Ewen Lescop. For the sake of completeness, and to demonstrate the direct extension of the concepts introduced in section 5.1, I nevertheless include it into this thesis.

2 Fast acquisition methods for multidimensional protein NMR

2.1 Resolution, Sensitivity and Time Requirements in NMR

The previous chapter has revealed that a wealth of information can be obtained from NMR spectroscopy about the structure, dynamics and kinetics of (bio-)molecules. This information is contained in the response of the nuclear spins placed in a static magnetic field to irradiation by radio-frequency electromagnetic waves.

This response can be obtained for many individual nuclear spins, and the accumulated information of all these spins contains information about molecular structure, dynamics and interactions. It is generally displayed in the form of a spectrum, i.e. a plot of signal intensity as a function of frequency.

The way from the spectrum to an atom-resolved molecular picture involves the quantification of the positions, heights and widths of spectral signals. The precision with which all these parameters can be extracted therefore determines the precision of the molecular picture that one gets in the end.

The two crucial factors that govern the precision with which the spectral parameters can be obtained turn out to be the *sensitivity* and the spectral *resolution*. The next sections will define these quantities.

2.1.1 Brief introduction of some basic concepts

When an ensemble of spins $\frac{1}{2}$ is placed in a magnetic field of strength B_0, the two eigenstates, denoted α and β, have different energies and the energy difference, $\Delta E = E_\beta - E_\alpha$, is given by the strength of the magnetic field and the gyromagnetic

ratio of the nucleus, γ:

$$\Delta E = \hbar\gamma B_0 \quad \text{or in frequency units } \Delta E = \hbar\omega_0 \tag{2.1}$$

Here $\hbar$ is Planck's constant devided by 2π and ω_0 is the Larmor frequency (in units of rad/s).

Because of this energy difference the eigenstates are not equally populated, but the populations, n_α, n_β, are given by the Boltzmann distribution:

$$n_{\alpha,eq} = \frac{1}{Z} N \exp(-E_\alpha/kT) \quad n_{\beta,eq} = \frac{1}{Z} N \exp(-E_\beta/kT) \tag{2.2}$$

where N is the total number of spins, k is Boltzmann's constant, and T is the temperature. Z denotes the partition function, and for a spin-1/2 it is given as $\exp(-E_\alpha/kt) + \exp(-E_\beta/kt)$.

At ambient temperature the energy difference ΔE is very small compared to kT (high temperature limit), and one can simplify equation 2.2 using $\exp(-x) \approx 1 - x$ and $Z \approx 2$ to

$$n_{\alpha,eq} = \frac{1}{2} N (1 - E_\alpha/kT) \quad n_{\beta,eq} = \frac{1}{2} N (1 - E_\beta/kT) \tag{2.3}$$

The population difference, also referred to as spin polarization, is very small at ambient temperature, even at the highest magnetic field strengths available today. For example, in an ensemble of ^{1}H spins, the population difference at B_0=1 T is only on the order of 10^{-4}.

2.1.2 Sensitivity in NMR spectroscopy

Sensitivity in NMR, as in any other physical method, describes the experimental ratio of signal to noise per unit time

$$\text{Sensitivity} = \frac{\text{Signal}}{\text{Noise}} \cdot \frac{1}{\text{unit experimental time}} \quad \text{or briefly} \quad \left(\frac{S}{N}\right)_t \tag{2.4}$$

Several factors influence the sensitivity. The signal S is proportional to the bulk magnetization, which is the product of the gyromagnetic ratio, γ, and the polarization as introduced above. The small population difference between nuclear spin eigenstates translates to a low sensitivity of NMR experiments, which has been a major concern of NMR from the early days on.

Other factors related to the sensitivity can be summarized as follows:

$$\left(\frac{S}{N}\right)_t \propto \frac{N_{spins}}{V} \, \gamma_{exc} \, {\gamma_{det}}^{3/2} \, {B_0}^{3/2} \, {n_s}^{1/2} \, \frac{1}{\sqrt{R_s(T_a \, T_s) + R_c(T_a \, T_c)}} \tag{2.5}$$

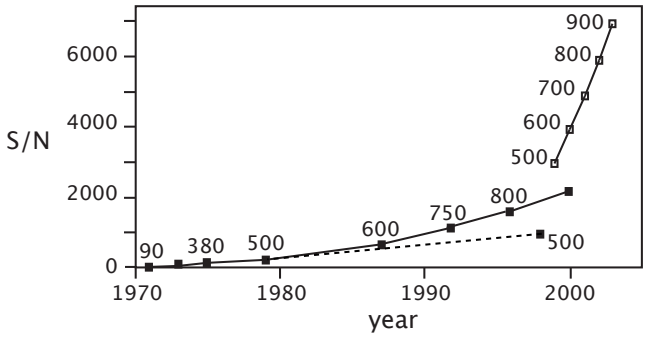


Figure 2.1 Evolution of the signal-to-noise ratio of 0.1% ethylbenzene in $CDCl_3$ for ^1H observe probes since 1970. The numbers above the black squares denote the magnetic field strength (^1H frequency in MHz). Open squares represent data measured with cryogenically cooled probes. The dashed line indicates the development of conventional probes for a 500 MHz spectrometer. Data adapted from Kovacs et al. (2005)

where T denotes the temperature in the preamplifier (a), the sample (s) and the radio-frequency coil (c). R_c is the resistance of the coil originating from an inductive coupling between the sample and the coil, whereas R_s is the resistance of the sample related to its conductivity. γ_{exc} and γ_{det} are the gyromagnetic ratios of the excited and detected spins, respectively, and N_{spins}/V is the number of spins per volume.

In many cases, data accumulation is necessary to obtain acceptable signal-to-noise. Adding up n_s scans increases the signal as n_s; the noise also increases, as $\sqrt{n_s}$, which gives an overall $\sqrt{n_s}$ dependence for the sensitivity.

Increasing the sensitivity of NMR has been and is still a very active field of development, and equation 2.5 indicates viable routes.

- The density of the spins (N_{spins}/V) can be increased to some extent. However, for biomolecules concentration is somewhat limited, and concentrations on the order of a few millimoles per litre can rarely be exceeded.

- High-γ nuclei are predicted to yield superior sensitivity. In protein NMR experiments (in liquid state) most experiments therefore excite and detect ^1H spins.

Various hardware developments have further increased sensitivity:

- The magnetic field strength B_0 has been increased by constant hardware development to about 21 T to date.

- Reduction of the noise level by lowering the temperature of the preamplifier as well as the coil by cryogenically cooled devices has been shown to yield sensitivity gains on the order of a factor of 2-4.

The hardware developments (probes, field strengths, electronics) have contributed to a great augmentation of signal-to-noise and this increase over the last decades is summarized in figure 2.1. This progress translates to a much higher precision of NMR data, collected in the same amout of time, or significantly reduced data accumulation times for achieving the same signal-to-noise.

In addition, there is very active research in the field of spin hyperpolarization, aiming at increasing the spin polarization from the usual 10^{-4}-10^{-5} to values that approach unity (Albert et al. (1994); Bowers and Weitekamp (1986); Muller-Warmuth and Meise-Gresch (1983))

2.1.3 Spectral resolution

Spectral resolution denotes the minimum frequency difference between two signals that can be identified as individual resonances in a spectrum. The resolution of an NMR spectrum is determined by the number of resonance lines present in a given frequency range (the peak density) and their linewidths.

The Larmor frequency of a nucleus depends on its gyromagnetic ratio and the magnetic field strength (see equation 2.1); in addition, the chemical environment of the nucleus alters the magnetic field strength that the nucleus actually senses. This quantity is expressed as chemical shift δ (in units of parts-per-milion, ppm). It is defined as a normalized frequency difference with respect to a reference compound:

$$\delta = \frac{\omega - \omega_{ref}}{\omega_{ref}} \cdot 10^6 \tag{2.6}$$

The resonance frequency then depends on δ, γ and B_0 as:

$$\omega = -\gamma \cdot (1 + \delta) \cdot B_0 \tag{2.7}$$

Because of this normalization with respect to a reference frequency, the chemical shift of a nucleus at a given position in a molecule is independent of the spectrometer operating frequency and depends only on physico-chemical properties of the molecule.

Another quantity that is related to resolution is the *digital resolution*, Δf, which is the width of the spectrum, sw relative to the number of data points in a spectrum (n):

$$\Delta f = \frac{sw}{n} \tag{2.8}$$

2.1.4 Fourier transform NMR

In early days of NMR, experiments were performed in a continuous-wave (CW) manner: the sample was irradiated with a continuous radiofrequency (rf) field placed in a magnetic field of which the field strength was varied. In the course of this variation of the field strength, also called sweeping, the nuclear spins come into resonance only when the radiofrequency matches the energy difference between the nuclear spin eigenstates.

Consider for example a sample containing molecules that have three different ^{1}H sites with different chemical shifts (thus three signals in frequency space). Sweeping the magnetic field strength while keeping the rf field constant brings the three different sites into resonance at different stages of the sweeping process.

This is akin to slit-techniques in optical spectroscopy, where the frequency spectrum is scanned by observing only a small part of the spectrum which is scanned over the whole frequency range. Because these techniques only observe a small portion of the frequency space at a time, they are sometimes called *single-channel* techniques.

Focussing at only one such slit of the spectrum at a time has several disadvantages. The sensitivity is low, because only a small slit is observed while a large fraction of the spectrum is neglected. In addition, achieving a high digital resolution Δf requires to detect a large number of points in the frequency range, N (see equation 2.8), and therefore requires long experimental times related to the slow sweeping of the frequency range.

The introduction of Fourier transform (FT) NMR (Ernst and Anderson (1966)) has today almost entirely replaced CW NMR and constituted a leap in NMR's applicability, resolution and sensitivity. In contrast to CW techniques, FT NMR excites the whole spetral range at the same time by short radiofrequency pulses, and the evolution of the excited spins is followed as a function of time. The distribution of intensities in frequency space, $I(\omega)$ is then obtained by a Fourier transformation of the signal detected as a function of time, $s(t)$, which is sampled from $t = 0$ to $t = t^{max}$.

$$I(\omega) = \int_0^{t^{max}} s(t) \exp(-i\omega t) \mathrm{d}t \tag{2.9}$$

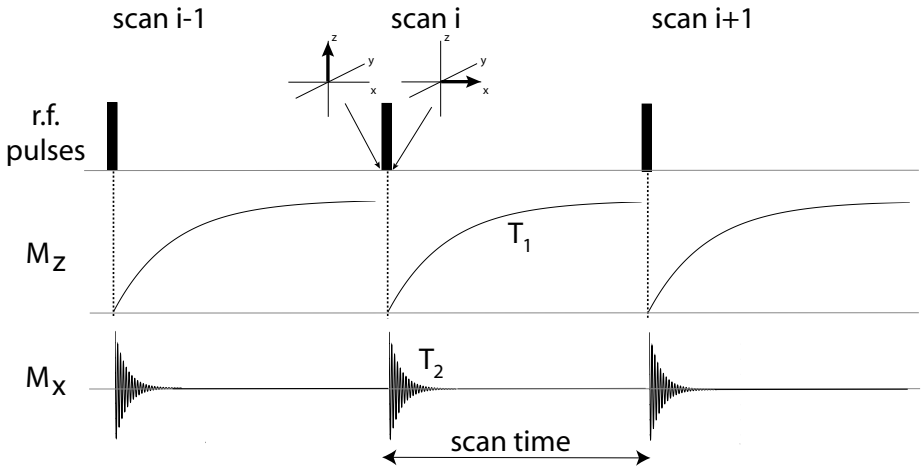


Figure 2.2 Schematic representation of 1D FT NMR. The evolution of the longitudinal component of the bulk magnetization, M_z, and the transverse component, M_x, are displayed during the three scans, each consisting of an excitation with a 90° pulse, subsequent detection of the transverse magnetization and a longitudinal relaxation delay. The insets represent pictorially the bulk magnetization vector (shown here only for one scan). The signal acquired in the different scans can be added for increased signal-to-noise before final Fourier transformation.

In practice, the sampling cannot be performed in a continuous manner, but the time-domain signal is sampled at n discrete steps separated by a so-called dwell time Δt. The integral transform in equation 2.9 therefore has to be replaced by a discrete Fourier transform.

The time-domain signal originates from the superposition of *all* the excited nuclei present, and Fourier transform NMR, unlike CW NMR, can be regarded as a *multi-channel* approach. This advantage, that all frequencies are sampled at the same time by observing a single continous signal in the time domain is referred to as *multiplex advantage*.

Figure 2.2 shows schematically the experimental realization of a simple one-dimensional (1D) FT NMR experiment. Although a rigorous treatment requires a quantum-mechanical description, a simple vector model introduced by Felix Bloch suffices for the moment to introduce some concepts, and it will be used at several places in this work.

In this classical picture, the bulk magnetization of the sample at thermal equilibrium is represented by a vector parallel to the external magnetic field (figure 2.2). The effect of a rf pulse (depicted by a narrow filled rectangle) can be represented in this picture by simple rotations of this vector. For example, a so-called 90° pulse results in the alignment of the bulk magnetization in the xy-plane of the reference

frame. This is referred to as *transverse magnetization*. The transverse magnetization precesses around the magnetic field with an angular frequency ω, inducing a current in the surrounding receiver coil. This is exactly the signal $s(t)$ that can be Fourier transformed according to equation 2.9.

The evolution of the longitudinal and transverse components during this simple one-pulse FT NMR experiment are schematically summarized in figure 2.2. For the sake of generality figure 2.2 assumes that the pulse sequence is repeated numerous times (e.g. in order to increase the signal-to-noise ratio) and three repetitions, or scans are shown. The transverse magnetization present after the pulse oscillates with the resonance frequency of the nucleus and this oscillation corresponds to a damped (decaying) sinusoidal function. The damping is due to a dephasing of the individual nuclear magnetic moments in the sample (a loss of coherence). This process is called transverse relaxation, denoted by a time constant T_2. Due to this decay the time-domain signal is called free-induction decay , or FID. (Note that the observed signal is in general the superposition of many such damped sinusoudals, according to the number of spins with different resonance frequency.)

The longitudinal component in this picture, M_z, is zero after a 90° pulse and it returns to its thermal equilibrium value via a process called longitudinal relaxation. This spin magnetization recovery is often assumed to be monoexponential with an associated time constant T_1.[1]

The transverse relaxation time constant is always smaller or maximally equal to T_1, and in the case of proteins typical T_1 and T_2 values for protons are on the order of 1 second and tens of milliseconds, respectively.

Some of the properties of discrete FT NMR will be needed in the following and are briefly introduced here.

When data are sampled discretely at evenly spaced intervals in time (separated by Δt), the highest frequency that can be unambiguously quantified is given by the Nyquist frequency:

$$f_n = \frac{1}{2\Delta t} \tag{2.10}$$

Frequencies that are higher than f_n are not suffiently sampled and appear in FT NMR as lower than they actually are (see figure 2.3). The frequency range that can be observed (i.e. the spectral width sw) is the interval $-f_n < \nu < f_n$ and thus depends on the sampling interval as

[1] Section 2.3 will adress this issue in more detail.

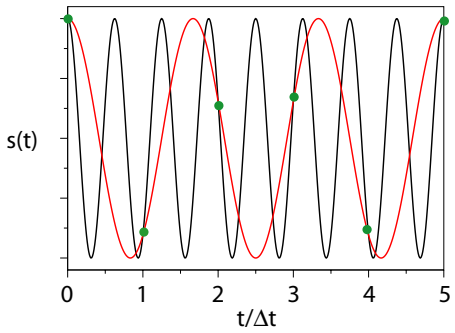


Figure 2.3 Nyquist theorem. A cosine function $\cos(2\pi\nu_0 \cdot \Delta t)$ with ν_0=1.6 (black curve) is sampled at discrete points separated by Δt (green filled circles). This sampling (less than twice per period) makes it appear at a frequency $\nu_a = \nu_0 - 2 \cdot f_n$ = 0.6 (red curve).

$$sw = \frac{1}{\Delta t} = 2f_n \tag{2.11}$$

The digital resolution Δf=sw/n is determined by the maximum data sampling time t^{max} as:

$$\Delta f = \frac{1}{t^{max}} = \frac{1}{n\ \Delta t} \tag{2.12}$$

This equation states that in order to achieve a certain digital resolution Δf, a number $n = sw/\Delta f$ of points has to be sampled along the time axis.

Another quantity that determines resolution is the width of the signals in frequency space, L. It depends on the decay rate of the corresponding signal in time space, R_2=$1/T_2$.

$$L = \frac{R_2}{\pi} \tag{2.13}$$

As the transverse relaxation rate of a given nucleus can be considered an intrinsic property under the experimental conditions, the line width L is also called *natural* line width. The natural line width can be considered as an upper limit of the resolution that one can achieve by FT NMR (or any other NMR technique). The line width can be related to the digital resolution (equation 2.12) as

$$\Delta f = \frac{1}{t^{max}} = L\left(\frac{\pi R_2^{-1}}{t^{max}}\right) \tag{2.14}$$

Optimal resolution is achieved when the digital resolution is on the order of the natural line width. Higher digital resolution will not further improve the ca-

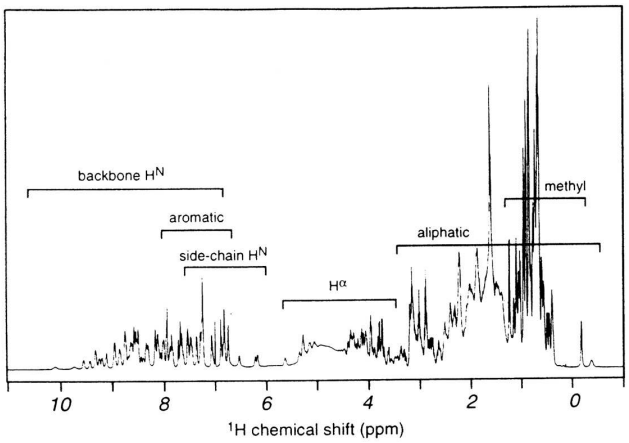


Figure 2.4 One-dimensional ^{1}H spectrum of the protein ubiquitin recorded at 500 MHz ^{1}H Larmor frequency. The frequency ranges of different types of ^{1}H spins are indicated.

pability to discern separate signals. Equation 2.14 predicts that optimal resolution is achieved setting $t^{max} = \pi R_2^{-1}$ or $n = (\pi R_2^{-1})/(\Delta t)$.

2.1.5 FT NMR applied to biological macromolecules: the need for more resolution

Figure 2.4 shows a ^{1}H spectrum of a small protein; strong overlap between signals corresponding to different sites in the molecule is observed, and only very few sites have resolved signals in this ^{1}H spectrum, which makes extraction of atom-resolved information almost impossible. Increasing the resolution is therefore a critical issue for the study of biomolecules by NMR and different approaches have been proposed.

The resolution is related to the peak density and the line widths of the individual signals. In order to achieve high resolution, the peak density should be low and the line widths should be as small as possible.

To keep the peak density low, a first prerequisite is to avoid the presence of molecules that are not of interest; biochemical purification is therefore indispensable for biomolecular NMR studies.

However, this is in general not sufficient: the intrinsic peak density in proteins and nucleic acids is high. This is related to the fact that these molecules are composed of repetitions of a relatively small number of building blocks that all share a common motif. As the chemical shift is mainly determined by local effects the

signal positions are confined within a relatively narrow range resulting in a high peak density even for small proteins.

Equation 2.7 shows that increasing the magnetic field strength B_0 is a way to increase resolution. Assuming that the line width (in Hertz) does not change with magnetic field strength, resolution increases linearly with B_0, which is one of the factors motivating the ongoing efforts in the hardware development.

The peak density can also be reduced by reducing the number of signals in the spectrum. This can be done by replacing NMR-active nuclei (such as ^{1}H) by NMR-invisible nuclei (such as ^{2}H). Deuteration with selective protonation of a few sites can be a way in this direction.

Besides the peak density, the width of the signal also determines the resolution. When factors related to the acquisition procedure are set optimal (high digital resolution, homogeneous magnetic field), it is the natural line width that determines the minimal line width. The natural line width is related to a number of factors, including the nature of the nucleus observed, the molecular environment of this nucleus (which relates to relaxation mechanisms available and their strength), the spin state of neighboring nuclei and the dynamics sensed by the given nucleus (which modulates the coupling strength to the available coupling mechanisms). For a given nucleus a reduction of the natural line width can therefore be achieved by either altering the dynamics (higher temperature, lower viscosity (Wand et al. (1998)), removing relaxation mechanisms by replacing proton spins by deuterons (Gardner and Kay (1998)), or by use of transverse relaxation optimized spectroscopic techniques (Fiala et al. (2000); Pervushin et al. (1997); Sklenar et al. (1993)).

However, all these approaches, from higher field strengths and biochemical purification to techniques that decrease the natural line width, prove insufficient to resolve a large number of individual nuclear sites. Detailed atom-resolved information about biomolecules is generally inaccessible to 1D NMR.

2.1.6 Multidimensional NMR

Jeener and later Ernst (Aue et al. (1976); Jeener (1971)) have proposed a very elegant way to increase resolution and, additionally, spectral information content, that has revolutionized NMR and opened the way for almost all modern applications: multidimensional Fourier Transform NMR (in the following referred to as nD FT NMR). Unlike 1D NMR, it searches to correlate resonance frequencies and thus spread signals from a one-dimensional into an n-dimensional frequency space. Not only does this spreading increase resolution, but the correlation of fre-

quencies along independent frequency axes reveals the presence of an interaction between spins resonating at the frequencies that are correlated. This correlation information is extremely valuable and gives insight into structure and dynamics and can also be used for spatial location in imaging experiments. Much of NMR's power is based on multidimensional techniques, and the range of applications makes NMR today probably one of the most versatile techniques, indispensable in fields as diverse as medical diagnostics and the investigation of molecular structures at atomic resolution.

2.1.6.1 The concept of nD NMR

The idea behind nD FT NMR, very much akin to 1D FT NMR, is to collect a signal that is a function of n time variables, $s(t_1, t_2, ..., t_n)$, and to reconstruct the complete n-dimensional spectral distribution, $I(\omega_1, \omega_2, ..., \omega_n)$, by Fourier transformation. For the case of a 2D signal this would be:

$$I(\omega_1, \omega_2) = \int_0^{t_1^{max}} \int_0^{t_2^{max}} s(t_1, t_2) \exp(-i\omega_1 t_1)\ \exp(-i\omega_2 t_2)\, \mathrm{d}t_1\, \mathrm{d}t_2 \tag{2.15}$$

From a formal point of view nD FT NMR thus appears as a natural extension of 1D NMR as can readily be seen by comparing equation 2.15 to equation 2.9. However, the formal analogy is misleading: Whereas a time-domain signal in one dimension, $s(t)$, can be acquired by physically observing the current induced in a detection coil by the spins' precession, one cannot sample the spins' evolution along two time axes simultaneously: time is a one-dimensional quantity by nature. Acquisition of a signal $s(t_1, t_2)$, is therefore done by repeating the basic experiment multiple times, detecting each time the full time domain signal in one dimension, and incrementing with each repetition of the experiment a delay parameter within the pulse sequence that reflects the time evolution of the remaining time domain. For the case of a 2D data set, this is realized with this general scheme:

$$\text{preparation} - \text{evolution in } t_1(\text{incremented}) - \text{mixing} - \text{acquisition } (t_2) \tag{2.16}$$

In this scheme, an initial preparation step excites the spins of interest, which are allowed to evolve according to their resonance frequency Ω_1 during an (incremented) delay t_1. At the end of this delay a mixing period transfers the magnetization to other sites - provided that there is an interaction mechanism allowing this transfer. The resonance frequency of the spins on which the magnetization resides after mixing, Ω_2, is then observed during the final acquisition period. The

final signal is thus a function of the two time variables that yields the desired 2D spectrum after 2D FT (equation 2.15). The presence of a signal at frequencies Ω_1 and Ω_2 in the 2D spectrum reveals the presence of the interaction mechanism between the two nuclei resonanting at Ω_1 and Ω_2.

In this scheme, the sampling of ni data points in the indirect time domain requires ni repetitions of the pulse sequence.

Extension to higher dimensionality is straightforward: additional mixing and evolution steps for each additional spectral dimension have to be introduced and all the data points in the indirect dimension have to be sampled independently.

Figure 2.5 shows an example of how multidimensional NMR techniques are used to increase resolution and information content. The sampling scheme along the indirect time domains, t_1, t_2 is also shown.

2.1.6.2 Resolution in nD FT NMR

The resolution enhancement brought about by nD NMR is explained by the high-dimensional space in which the signals are distributed.

For the width of the resonance lines similar considerations apply as in 1D NMR. Just as in the direct frequency domain, the natural line width is determined by the transverse relaxation of the nucleus that evolves during the corresponding evolution time. The natural line width, given by equation 2.13 on page 40, can be considered an upper bound of the achievable resolution.

The *digital* resolution, just as in 1D, depends on the maximum delay that is sampled, t^{max}. For equidistant sampling of ni data points separated by Δt, it is given by

$$\Delta f = \frac{1}{t^{max}} = \frac{1}{ni\ \Delta t} = L\left(\frac{\pi R_2^{-1}}{t^{max}}\right) \tag{2.17}$$

As before, optimal digital resolution is achieved setting $t^{max} = \pi R_2^{-1}$ or $ni = (\pi R_2^{-1})/(\Delta t)$. Unlike 1D NMR, achieving higher resolution (while keeping the spectral width sw constant), can only be achieved by collecting more repetitions of the experiment.

To put these equations in a context and demonstrate what the numbers are in a typical case, consider the sampling of the ^{15}N time domain in a magnetic field of 14.1 T (600 MHz ^{1}H frequency). Assume that the spectral width to be covered is 35 ppm (corresponding to 2160 Hz) and further assume a transverse relaxation time constant R_2^{-1} of 70 ms , which is a typical value found in a small to medium-sized protein (Rovnyak et al. (2004b)). In this case, if a resolution is desired that allows to distinguish signals separated by twice the natural line width ($ni = (\pi R_2^{-1})/(2\Delta t)$)

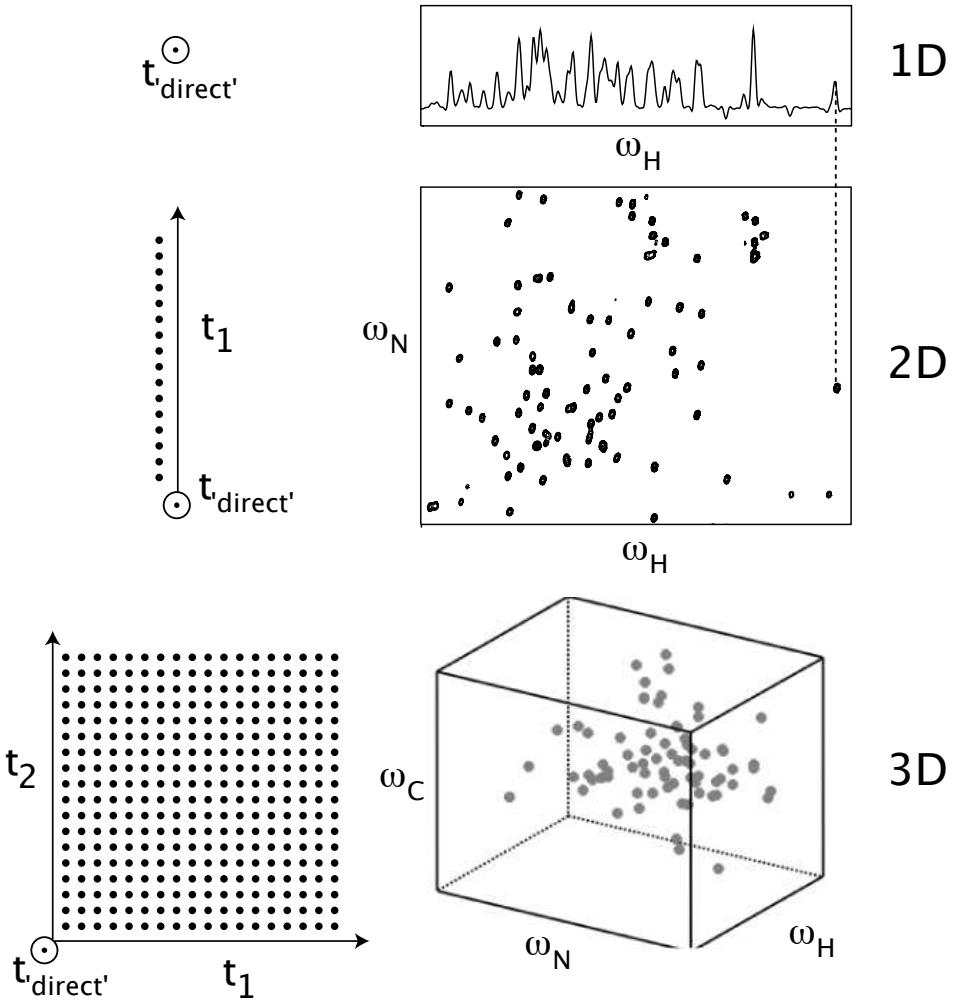


Figure 2.5 Increase in resolution and information content achieved by multidimensional NMR, demonstrated with the example of the protein ubiquitin. The 1D spectrum of the amide proton region is displayed in the upper panel (right). It was obtained by Fourier transformation of time domain data acquired along an axis t_{direct} (pointing out of the plane). The central panel shows a correlation spectrum of the same amide protons to the attached ^{15}N nuclei. The time-domain data set is schematically indicated on the left side. Each point corresponds to the acquisition of a free-induction decay along the direct time domain. Lower panel: A 3D spectrum correlating the amide proton, its attached nitrogen and the neighboring carbonyl atom. Two independent time domains have to be sampled, as indicated by the sampling grid on the left. Note how the peaks that are overlapping in the 1D spectrum are resolved in the 2D and/or 3D spectra. The experimental time required for the recording of these data sets is proportional to the number of points on the sampling grid.

more than 230 data samples have to be acquired. This is found one of the major drawbacks of multidimensional NMR: Hundreds of repetitions of the experiment with incrementation of evolution parameters have to be performed to sample the additional time domain(s).

2.1.6.3 nD FT NMR and the multiplex advantage

At first sight, the nD acquisition scheme seems to relinquish the advantages introduced into 1D NMR by the FT approach, namely that a full spectrum can be obtained in one single scan. So do multidimensional NMR techniques still benefit from the multiplex advantage afforded by 1D FT NMR ?

At this point, one should distinguish between two different issues, sensitivity and experimental time.

In terms of sensitivity, the full multiplex advantage is retained: all the resonance lines contribute to the signal detected in each and every scan, such that in a first approximation sensitivity is not reduced. Thus, to a first approximation, it does not make a difference if multiple scans are accumulated by repeating exactly the same sequence or if the same number of scans is used to sample an additional time domain by increasing in each scan a delay parameter within the pulse sequence. Multidimensional NMR therefore comes at (almost) no cost in terms of sensitivity and in cases where sensitivity is the limiting factor governing the experimental time it even comes at no cost in terms of experimental time.[2]

The real drawback of nD NMR as compared to its one-dimensional counterpart is the minimum number of scans (and thus the minimum experimental duration) required for the sampling of the full n-dimensional data set, even in cases where sensitivity is abundant in few scans. The above considerations have shown that for optimal spectral resolution hundreds of scans have to be sampled in each time domain. As the spin evolutions along different time domains are independent, the number of scans grows exponentially with the number of dimensions. The next section will take a closer look at the typical experimental times needed for the sampling of the data space.

[2]Here signal losses during the generally longer pulse sequences used for higher-dimensional experiments (including relaxation losses), and losses due to quadrature detection are not considered.

2.1.6.4 Time requirements of nD FT NMR

The minimum experimental time for the recording of an nD data set (with indirect time domains $t_1, \ldots, t_{n-1}$) is given by the duration of one scan, T_{scan}, times the number of scans that are performed:

$$T_{experiment} = T_{scan} \cdot \underbrace{\left(n_{repet}\, n_{pc}\, 2^{n-1}\, (ni_1\, ni_2\, \cdots\, ni_{n-1}) \right)}_{\text{number of scans}} \tag{2.18}$$

Here ni_k is the number of time points (increments) acquired in the k-th time domain (determined by the spectral width sw_k and the desired resolution in this dimension) and the factor 2^{n-1} arises from the quadrature detection scheme in each of the $n-1$ indirect dimensions, when using one of the usual procedures (States (States et al. (1982)), TPPI (Marion and Wuthrich (1983)) or TPPI-States (Marion et al. (1989)). n_{pc} is the number of repetitions of the experiment that has to be performed to complete phase cycling (which is used for artifact suppression). n_{repet} is the number of repetitions that is done for signal accumulation to increase sensitivity. Using modern hardware, sensitivity is often sufficient using n_{repet}=1, and in the situations discussed below this is always assumed; experimental times are thus not considered to be sensitivity-limited, but determined by data sampling. T_{scan} includes the duration of the rf pulse sequence (rf pulses, pulsed field gradients and delays), the acquisition period and the delay between consecutive scans (interscan delay; see figure 2.2).

When studying biological macromolecules, typical values for T_{scan} are on the order of one second. As for the number of scans, typical values are on the order of 50-100 per indirect time domain. With these values and assuming a phase cycling scheme with $n_{pc} = 2$ one obtains the following typical values for the minimum durations of multidimensional data acquisition:

1D	2D	3D	4D	5D
2 seconds	3 minutes	11 hours	92 days	600 months

Because of the long experimental times 4D and 5D spectra are very rare and only acquired when the number of sampled data points can be significantly reduced by some means. For the set of experiments that typically have to be performed for structural studies of biomolecules (several 3D spectra) the spectrometer time required is on the order of several days to weeks.

2.1.7 Speeding up multidimensional data acquisition

The long data acquisition times associated with the data sampling is found very confining for multiple reasons:

- Many biological samples tend to have a limited stability in time. Finishing the whole set of required experiments within the 'life time' of the sample is therefore crucial to obtain reliable information.
- High-throughput structural genomics projects require the collection of large amounts of data in short times. This also applies to large scale facilities, where sensitivity is good because of high magnetic fields, but instrumental time expensive.
- The study of more and more complex systems requires ever higher dimensional data sets ($\geq$ 4D), that would take unacceptably long experimental times using standard schemes.
- Following (bio-)chemical reactions in real time requires methods that are capable of accessing the time scales of these events while resolving individual nuclear sites. Many biologically relevant reactions occur on time scales of seconds (or less), which makes them often accessible to 1D techniques but excludes the use of site-resolved nD NMR studies using standard nD techniques.

In addition, unlike in the early days of NMR, where sensitivity was often the factor dominating the experimental duration, technical advances in the last decades have made available high sensitivity in times that are much shorter than the durations dictated by data sampling considerations, a situation that has been known as *sampling-limited regime*. The last years have therefore witnessed intense research efforts to eliminate the time barriers of nD NMR.

As indicated by equation 2.18, there are two possibilities for accelerating nD techniques: the reduction of the time required to sample one data point, and the reduction of the number of data points that have to be sampled. All the proposed methods are based on these two strategies and in many cases different concepts can be combined, offering now a large pool of techniques for fast nD NMR.

The remainder of this chapter deals with the concepts that have been proposed. In the first part (section 2.2), approaches for reducing the number of sampled data points are reviewed. Section 2.3 will then introduce methods that allow a significant reduction of the interscan delay while retaining high sensitivity. The main

focus will be on methods used in protein NMR spectroscopy that have been used and developed in the course of this thesis.

2.2 Accelerated nD NMR I: Reduced number of scans

Different approaches have been proposed to reduce the minimum number of scans that have to be performed. Here I will not discuss the class of methods that use non-linear data sampling in combination with non-linear processing methods, such as the filter diagonalization method (Chen et al. (2003); Mandelshtam (2000)), maximum entropy methods (Hoch and Stern (1996); Pons et al. (1996)), three-way decomposition (Orekhov et al. (2001, 2003)), Lagrange interpolation of non-equispaced data followed by regular FT (Marion (2005)) or brute-force FT of non-uniformly sampled data (Marion (2006); Pannetier et al. (2007)). Several review articles are available that cover these methods (Freeman and Kupce (2003); Kupce et al. (2003); Malmodin and Billeter (2005a)).

2.2.1 Phase cycling

Considering equation 2.18, reducing the number n_{pc} of phase-cycled scans is a viable approach for speeding up data acquisition. Phase cycling is done by repeating the experiment n_{pc} times with different settings of pulse/receiver phases. It is performed for suppression of spectral artefacts.

Hardware developments have allowed to reduce the number of steps in phase cycles while preserving good artefact suppression: First, increased spectrometer stability often allows a clean annulation of artefacts by summing only few (two, four...) phase-cycled scans. Second, the introduction of pulsed field gradients has revolutionized artefact suppression and coherence selection strategies. Rather than by extensive phase cycles, the desired coherence pathways can be selected in *one* single scan by the use of pulsed field gradients, and they are applied in almost all modern pulse sequences (see, e.g. Bax and Pochapsky (1992); Kontaxis et al. (1994)). In addition, spectroscopic methods have been proposed that allow to move axial peaks to regions of the spectrum that do not contain signals of interest (Simorre and Marion (1991)). To conclude, in many cases phase cycling for artefact suppression can be reduced or completely skipped.

When the number of phase-cycled scans is already reduced to a minimum, the next step is the reduction of data points sampled in the indirect time domains.

2.2.2 Spectral aliasing

Increasing the step size Δt between data samples results in aliasing (back-folding) of peaks that have frequencies outside the spectral width (equation 2.11). This can lead to accidental peak overlap and makes it difficult to know the real frequency of aliased peaks.

However, if peak positions are known before, one can optimize the spectral width such that the aliased peaks fall in empty regions of the spectrum or onto peaks that are not of interest. It is common practice to use spectral widths that are slightly smaller than the maximum frequency difference of peaks along this dimension.

A systematic approach can be envisaged: If the peak positions are known along indirect dimension(s), a numerical optimization of the spectral widths that minimize peak overlap can be performed. In this way, one can obtain very dense spectra with small spectral widths. The spectrum can then be unfolded with the knowledge of the real (not aliased) peak positions (Jeannerat (2007); Lescop et al. (2007).

This approach needs a prior knowledge of the spectrum and one may ask the question in which conditions it can be useful to record a spectrum in which the peak positions are already known. Two situations may be of interest:

In experiments where the peak *intensities* are of interest, and the positions are already known, a systematic spectral aliasing can speed up data acquisition while retaining all the desired information. This is the case when measuring a series of nD spectra and varying some parameter from one experiment to the other, e.g. a delay for the measurement of relaxation rates or diffusion constants.

In addition, extensive spectral aliasing can be useful for fast collection of high-dimensional ($\geq$ 3D) spectra that are extensions of lower-dimensional spectra already known. For example, think of a H-N-CO experiment: If the ^{1}H and ^{15}N peak positions are already known from a 2D spectrum, then a 3D HNCO data set can be recorded where the spectral width in the ^{15}N dimension is reduced to a minimum. A small number of data points along the folded ^{15}N dimension is then sufficient and only along the unknown ^{13}C domain the usual number of data points has to be recorded. This approach can then be extended from the known ^{1}H, ^{15}N and ^{13}CO peak positions to record 4D H-N-CO-CX experiments (with CX=CA or CB). How such a systematic spectral aliasing can be used is shown in the appendix, page 219).

2.2.3 Linear prediction

If the time domain is not sampled up to sufficiently long t^{max}, as required for optimal spectral resolution (see equation 2.14) one speaks of a truncated FID. This is general practice for keeping experimental times reasonably short. The consequence of such a truncation , besides decreased resolution, is a distorted line shape (sinc wiggles). Linear prediction (LP) allows to extrapolate experimental data points (Barkhuijsen et al. (1985)) and to extend the tail of a time-domain signal.

LP assumes that the time-domain signal consists of a superposition of complex sinusoids, and fits a number of such sinusoidal functions to the experimental data points. LP relies on a regular sampling grid.

LP is often used to extend the FID to longer t^{max}, and it may also be used to predict missing data points that are corrupted by instrumental failure (e.g. transmitter breakthrough).

2.2.4 Reduced dimensionality or Projection NMR

The speed problem of nD NMR comes from the need to sample several time domains *independently*. Reduced dimensionality techniques attack this problem by linking together the stepwise incrementation of different evolution periods (Brutscher et al. (1994); Simorre et al. (1994); Szyperski et al. (1993a,b)).

Consider a three-dimensional experiment, where two indirect evolution times t_1, t_2 are varied independently, and assume that an NMR signal has a frequency Ω_1 in t_1 and Ω_2 in t_2. At the end of the first evolution time, a signal $\cos(\Omega_1 t_1)$ is transferred to the second spin (with frequency Ω_2), and after the second evolution time the signal prior to acquisition is modulated as

$$\cos(\Omega_1 t_1)\cos(\Omega_2 t_2) \tag{2.19}$$

Quadrature detection of both time domains requires additional acquisition of the corresponding signals modulated as $\cos(\Omega_1 t_1)\sin(\Omega_2 t_2)$, $\sin(\Omega_1 t_1)\cos(\Omega_2 t_2)$ and $\sin(\Omega_1 t_1)\sin(\Omega_2 t_2)$.

In the reduced dimensionality approach (also known as projection NMR), the two evolution times are incremented jointly as a function of a single evolution time t, thus t_1 and t_2 are replaced by t and $t\lambda$, where $\lambda = \Delta t_2/\Delta t_1$ is a scaling factor between the two evolution periods. Figure 2.6 shows the data sampling along such

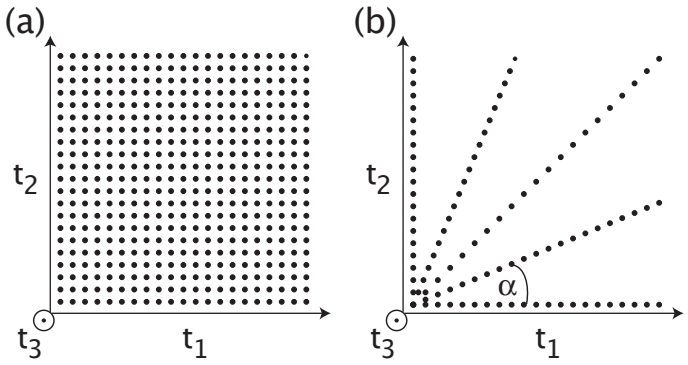


Figure 2.6 (a) Standard scheme for sampling of two indendent time domains (as introduced in figure 2.5). (b) Reduced dimensionality approach. The two evolution periods are linked together, as explained in the text. Data are acquired along an axis that is tilted with an angle α ($\lambda = \tan\alpha$) with respect to t_1.

tilted axes (corresponding to different values of λ).

In analogy to equation 2.19, this results in a modulation of the signal according to

$$\cos(\Omega_1 t)\cos(\lambda\,\Omega_2 t) \qquad \text{(S1)}$$

and as above, quadrature detection with respect to the two evolution periods can be performed, yielding additionally:

$$\sin(\Omega_1 t)\cos(\lambda\,\Omega_2 t) \qquad \text{(S2)}$$
$$\cos(\Omega_1 t)\sin(\lambda\,\Omega_2 t) \qquad \text{(S3)}$$
$$\sin(\Omega_1 t)\sin(\lambda\,\Omega_2 t) \qquad \text{(S4)}$$

Using trigonometric relations (S1) can be written as

$$\frac{1}{2}\{\cos(\Omega_1 t + \lambda\,\Omega_2 t) + \cos(\Omega_1 t - \lambda\,\Omega_2 t)\} \qquad \text{(S1')}$$

and similar relations hold for (S2) to (S4). Inspection of (S1') reveals that the acquisition along tilted axes results in the detection of *weighted sums and differences* of chemical shifts Ω_1 and Ω_2, rather than individual shifts.

Different possibilities have been proposed to analyse data acquired in this way. In the original approach (Brutscher et al. (1994); Simorre et al. (1994); Szyperski et al. (1993a,b)), quadrature detection was performed for only one of the two evolution periods, thus only (S1) and (S2) were recorded. The resulting spectrum contains a chemical shift doublet with peaks at $\Omega_1 + \lambda\Omega_2$ and $\Omega_1 - \lambda\Omega_2$. Determining their position allows to identify Ω_1 and Ω_2.

If quadrature detection is done for both nuclei (i.e. (S3) and (S4) are recorded additionally), then the two peaks at $\Omega_1 + \lambda\Omega_2$ and $\Omega_1 - \lambda\Omega_2$ can be separated in two subspectra (Brutscher et al. (1995b)).

This approach has been extended to higher dimensionality and formalized as G-matrix Fourier Transform (GFT) NMR (Kim and Szyperski (2003)). In principle, the reduced dimensionality approach can be applied twice for two (or more) distinct sets of chemical shifts, where the chemical shifts of each set are sampled in a joint manner (Atreya et al. (2005)).

The major advantage of projection NMR is speed: The number of data points that have to be sampled is considerably reduced with respect to the standard scheme (compare figures 2.6 (a) and (b)), at least as long as a small number of tilted planes is sufficient to obtain the desired information (small to medium-sized proteins). This makes it possible to record high dimensional (4D, 5D,...) information, which is inaccessible to the standard scheme, by reduction to low dimensionality.

In terms of sensitivity, reducing K frequency domains to a (K-1)-dimensional spectrum results in a sensitivity loss of a factor of $\sqrt{2}$ with respect to the corresponding K-dimensional spectrum. This loss can be partially compensated by exploiting spectral symmetry relations (Brutscher et al. (1995a)).

Because of the reduction of the dimensionality of the spectrum, accidental overlap is more likely than in the corresponding higher-dimensional spectrum. The reduced dimensionality approach is therefore especially interesting when the spectral space is relatively sparse, e.g. in protein resonance assignment experiments, but it is not limited to these experiments.

2.2.4.1 Projection reconstruction technique

The way that data are recorded in the reduced dimensionality approach can also be viewed in a different light (Kupce and Freeman (2003b, 2004a)).

Data sampling in reduced dimensionality experiments is done along a tilted axis in the multidimensional time space (see figure 2.6(b)). A Fourier theorem states that a section through the origin of a two-dimensional time-domain signal $s(t_1, t_2)$, which is inclined at an angle α with respect to the t_1 axis, transforms as the projection onto a line through the origin of the two-dimensional frequency-domain signal $S(\omega_1, \omega_2)$, with the same angle α with respect to the ω_1 axis (Nagayama et al. (1978)).

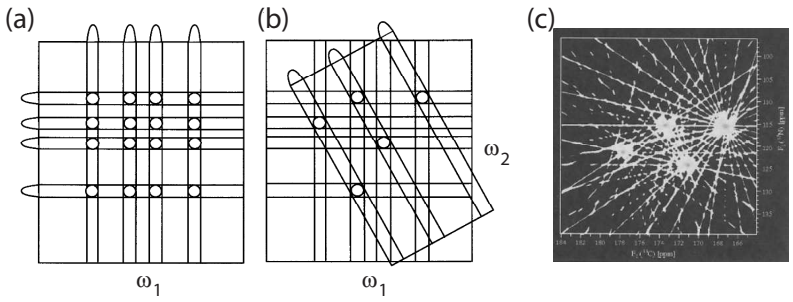


Figure 2.7 Projection reconstruction technique. (a) When a two-dimensional spectrum is reconstructed from two one-dimensional projections with 4 peaks in each projection, then 16 potential cross-peaks are possible. The problem is clearly underdetermined with only two projections and up to 16-4 of theses peaks may not be real. (b) The degeneracy is released when recording another projection at a different angle, and in this case the peak positions of the five crosspeaks are unambiguously determined from the three projections. However, in many cases more projections have to be recorded. (c) Experimental example of a typical plane of a H-N-CO spectrum of a protein containing 4 peaks, which is reconstructed from 18 projections evenly distributed from -90° to + 90°. Note that this reconstruction algorithm (sum-algorithm) results in noise ridges that are spread from the peak positions in the directions of the projections. These figures are reproduced from Kupce and Freeman (2004b).

In other words: The 2D spectra ((Ω_1, Ω_2),Ω_3), that results from Fourier transformation of the data set acquired along the tilted axis in the (t_1, t_2) plane can be seen as *projections* of the 3D $(\Omega_1, \Omega_2, \Omega_3)$ spectrum onto two dimensions. The projection angle α is given as $\tan\alpha = \Delta t_1/\Delta t_2$.

Lower-dimensional projections of an object are found in many other fields, such as astronomy, electron microscopy, computer tomography and other imaging techniques, and in these fields it seems obvious to *reconstruct* the real three-dimensional object from the two-dimensional projections. This operation is done by an inverse Radon transform, empirical backprojection schemes or iterative procedures.

In analogy, reconstructing a higher-dimensional spectrum from lower dimensional projections has also been proposed for NMR spectra (Coggins et al. (2005); Kupce and Freeman (2003b, 2004a,b); Venters et al. (2005)). Figure 2.7 (c) illustrates the sum-algorithm , or backprojection algorithm, where the intensities of all projections are summed up. Note that this reconstruction algorithm introduces projection artifacts, that manifest as ridges along the directions of the projections. Another algorithm, that avoids these artifacts is the lower-value approach. Rather than summing up all projections, it compares at each point of the spectrum the intensities of the different projections, and uses their lowest value. While removing the projection ridges, the resulting sensitivity is reduced: The reconstructed signal

corresponds to the signal of only one projection. Recording more projections thus does not increase the sensitivity. A hybrid algorithm between these two has also been proposed (Venters et al. (2005)).

The main advantage of the reconstruction techniques of projected (reduced dimensionality) spectra with respect to the original GFT approach is, that the final result is a usual high-dimensional spectrum. Most NMR spectroscopists are used to looking at spectra, rather than lists of frequencies, which is the outcome from the GFT approach described above.

The reduced dimensionality approach can be viewed from yet another angle: The recording of data points along tilted axes can be viewed as replacing the Cartesian sampling grid by a polar sampling grid with a radial and an angular component. In such a polar coordinate system, one may implement a straightforward Fourier transform to create directly the frequency domain spectrum along the two projected dimensions (Marion (2006)).

2.2.5 Hadamard NMR spectroscopy

The large number of NMR-active spins, and the corresponding large number of resonance lines, is at the origin of the strong overlap in low-dimensional spectra. Besides increasing the spectral dimensionality, resolution could also be enhanced by labeling specifically only few sites in the molecule with NMR-visible nuclei, which results in a simplified, and potentially resolved spectrum. Imagine, for example a protein that is protonated only at glycine residues, whereas all other amino acids are deuterated. Even a simple 1D ^{1}H spectrum would probably allow to resolve many individual glycine protons in such a diluted spectrum. In practice, this strategy is generally not applied: the high cost of such selective biochemical labeling, and the need for several samples with different labeling schemes, makes it unattractive.

Hadamard NMR can be seen as a spectroscopic alternative to such biochemical labeling. Rather than selecting specific sites by biochemical methods, it makes use of radiofrequency pulses that are applied selectively to a number of frequencies. These selective pulses replace the usual free evolution periods in a multidimensional pulse sequence. Consider for example a ^{1}H-^{15}N correlation experiment. Replacing the t_1 evolution period by an appropriate selective ^{15}N pulse would result in a 1D ^{1}H spectrum, which contains only signals from those ^{1}H spins that are coupled to a ^{15}N nucleus with a resonance frequency that matches the frequency of the selective pulse. In other words, the obtained 1D ^{1}H spectrum can be seen as a

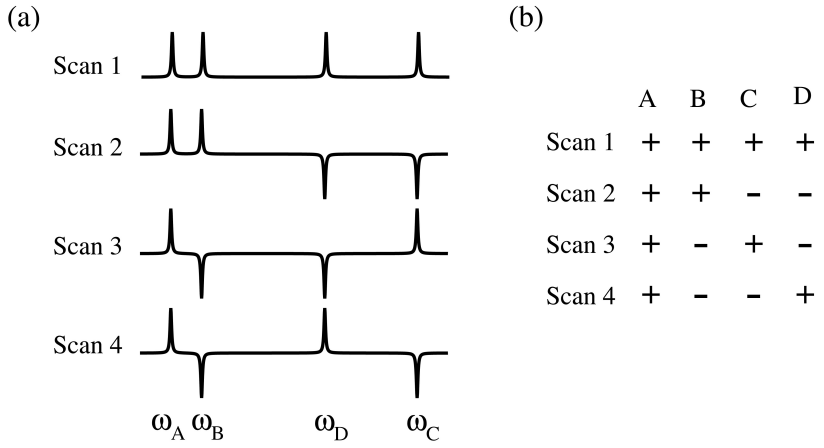


Figure 2.8 Hadamard encoded NMR spectroscopy shown for the example of a H-4 matrix. Four lines, A-D, are manipulated by selective pulses so, that their signs are positive or negative (a). From these four scans, the spectra containing only one line per spectrum (A, B, C, D) can be retrieved by summing the four scans according to the H-4 matrix (b).

slice of the 2D spectrum at the chosen ^{15}N frequency.

Although this approach would thus resolve certain sites in the molecule it has a drawback: In order to cover the *whole* frequency range, many repetitions would be necessary with different selective pulses, and each such repetition would contain the signal of only one slice. This frequency-space approach actually closely resembles the continuous-wave techniques of NMR's early days, with the associated drawbacks in sensitivity. Nevertheless, a slightly modified version of this idea, Hadamard spectroscopy, brings in again the multiplex advantage of FT NMR.

The trick of Hadamard NMR spectroscopy is the following: Rather than applying *one* selective pulse, an array of encoding pulses at different frequencies (different slices) is applied *simultaneously*. A number of such scans are performed, and in each scan the pulses are applied so, that the signal of different slices has either positive or negative sign, according to a so-called Hadamard-matrix (Hadamard (1893)).

Figure 2.8 shows this principle applied to the example of four selective encoding pulses. Four scans are performed and in each scan the signal of all slices is recorded. The individual peaks can then be retrieved by summing the scans according to the Hadamard matrix, figure 2.8(b). Note that completion of all scans of the Hadamard matrix is indispensable to extract individual spectra. Because in each scan the full signal of *all* slices is detected, the multiplex advantage is retained,

making it fundamentally different to the frequency-space CW techniques.

In addition to the shown four-by-four (H-4) matrix, Hadamard matrices of order 2^k and of order $4k$ can be defined, where k is an integer (H-2, H-4, H-8, H-12, etc).

Hadamard encoding instead of the usual time-encoding has a speed advantage whenever the number of Hadamard encoding pulses is smaller than the number of increments in an FT approach would be. Hadamard NMR spectroscopy is therefore especially interesting in two cases.

(i) If the spectral space in the indirect dimension contains only few peaks, or if only a few sites in the spectrum are of interest, then one can focus on these sites by manipulation with selective Hadamard encoding. In this context, focussing on an active site in a protein during a reaction or upon ligand binding (Feliz et al. (2006)), and the study of small molecules with only few lines (Kupce and Freeman (2003a)) have been proposed. In this case, the encoding pulses have in general a small bandwidth focussing on few sites; this techniques is called *site-selective* Hadamard approach.

(ii) If the peak density in the indirect dimension is high, and all peaks are of interest, then a site-selective approach is not attractive: to cover the whole spectral range, a large number of Hadamard pulses would have to be performed, resulting again in many mandatory scans and long experimental times. In this case a *band-selective* approach can be used: A rather small number of Hadamard encoding pulses with relatively large bandwidth can be more or less evenly distributed over the frequency range of interest. All the peaks resonating within the bandwidth of the individual pulses would thus not be resolved in the Hadamard-encoded dimension. However, the additional spectral dimension(s) may be sufficient to resolve the peaks within one band .

This band-selective approach was first applied to a protein HNCO experiment, where the CO frequency space is divided in four bands by selective encoding pulses. The correlation peaks are thus dispersed in four different 2D H-N edited spectra, where the H-N spin pairs found in different spectra have a neighboring carbonyl ^{13}C with a different chemical shift (Brutscher (2004)). A related approach allows to separate methyl correlation peaks into different 2D spectra according to the chemical shift of the neighboring ^{13}C atom, which depends largely on the type of the amino acid. In this way, separate spectra are obtained for different types of amino acids (Van Melckebeke et al. (2004)). In these two cases the Hadamard dimension serves as a third spectral dimension that allows to separate peaks in different 2D spectra, and only 4 data points have to be sampled for the additional dimension.

The speed advantage of these methods comes from the fact that only a few

scans have to be performed for the additional spectral dimension.

A drawback of the method is, that the chemical shifts in the additional frequency dimension have to be known. Setting up such an experiment thus requires some prior knowledge. However, a simple 1D spectrum of the frequency dimension that one wants to Hadamard-encode suffices and can generally be obtained very fast.

An additional drawback may arise when the encoding pulses are long. Relaxation losses may then translate to reduced sensitivity in Hadamard experiments compared to their FT NMR counterparts.

A Hadamard-encoded technique for very fast ($\sim$1 second) recording of site-resolved ^{1}H-^{15}N protein spectra is shown in section 7.1.

2.2.6 Single scan NMR

The long experimental time in standard nD FT NMR comes from the fact, that the increments in the indirect time domains are recorded in a *sequential* manner; i.e. in each scan *one* value of the incremented time period is obtained. In an ingenious approach proposed recently, this serial data acquisition is replaced by a *parallelized* scheme, where *different* evolution times are encoded in one single scan (Frydman et al. (2002)).

This single scan or ultrafast NMR approach imposes different evolution times on different parts of the sample. Although recently an alternative approach has been proposed (Bhattacharyya and Frydman (2006)), I will focus here on the original scheme, spatial encoding using field gradients. In this approach, applying a strong magnetic field gradient (typically along the sample axis) creates an inhomogeneous spectral distribution within the sample. The resonance frequency is then space-dependent and the sample can be viewed as composed of independent subsamples , which can be manipulated independently by frequency-selective rf pulses.

Now, a train of selective rf pulses, the frequency of which is incremented from one pulse to the other, or a frequency-chirped pulse can be applied. In this way spins located in different subsamples are on resonance with the pulses at different times and are allowed to evolve for different amounts of time. Therefore different subsamples will accumulate different phase (figure 2.9 (1)). The effects of the gradients on the accumulated phase are cancelled out and the phase ϕ of the magnetization at the conclusion of this encoding procedure is only a function of the

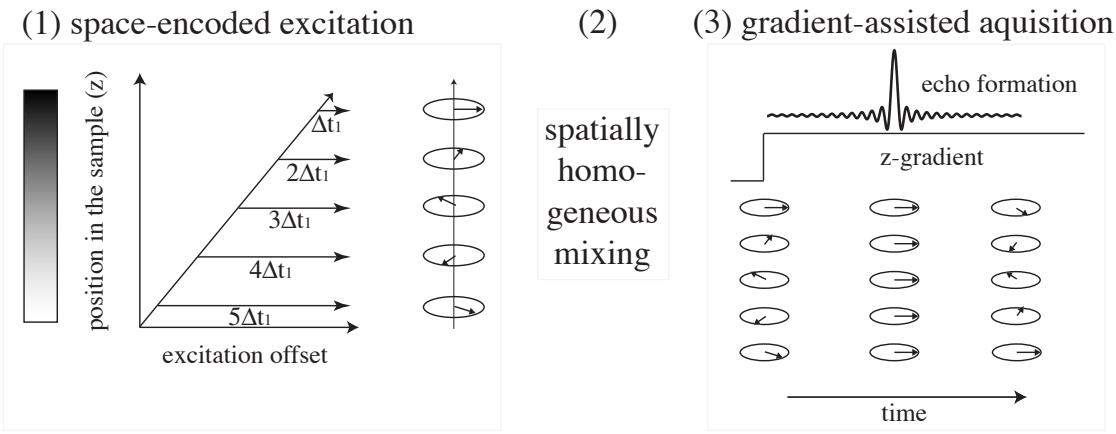


Figure 2.9 Principle of single scan (ultrafast) NMR. (1) Application of a linear z-gradient creates the spatial frequency dependence along the sample. Different subsamples (here depicted by 5 slices) can be manipulated by either a train of discrete selective pulses (Frydman et al. (2002)) or a continuous frequency-chirped pulse (Pelupessy (2003); Shrot et al. (2004)). This results in a helical winding of the magnetization located in different subsamples . (2) A spatially homogeneous mixing sequence preserves this space-dependent winding. (3) Acquisition in the presence of a gradient with identical geometry as the initial encoding-gradient unwinds the magnetization and leads to an echo (center), after a time determined by the chemical shift of the involved nuclei. The acquisition gradient is oscillated numberous times (with positive and negative sign) to monitor the evolution of the detected nucleus (direct dimension), and only the first positive gradient is shown (see also figure 2.10

spatial coordinate z and the chemical shift Ω_1:

$$\phi(z) = C\Omega_1(z - z_0) \tag{2.20}$$

with C a spatio-temporal ratio under the experimentalist's control.

After this spatial encoding a spatially homogeneous mixing sequence is applied, just as in standard experiments. The linear phase gradient[3] is preserved by this mixing sequence, yielding a helical winding of magnetization prior to acquisition (figure 2.9 (3)). Due to this winding, the net magnetization over the sample volume is zero.

Acquisition relies on a technique used in magnetic resonance imaging, called echo planar imaging (EPI): application of an acquisition gradient of strength G_a and of the same spatial dependence as the encoding gradient unwinds the magnetization. During this gradient, an echo will form when the following condition is

[3]This could also be a gradient of amplitudes rather than a phase gradient, depending on the experimental scheme used (Shrot et al. (2004))

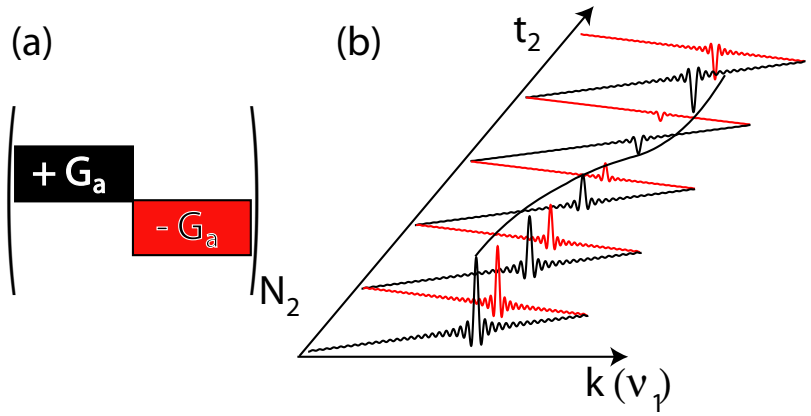


Figure 2.10 Data sampling in ultrafast NMR. During each oscillated gradient an echo (black: positive gradient, red: negative gradient) is formed at $k = -C\Omega_1$ (see figure 2.9) and the echoes are modulated in t_2 with Ω_2. Fourier transformation along t_2 reveals Ω_2.

fulfilled:

$$k = \int_0^{t_2} \gamma_a G_a(t) \mathrm{d}t = -C\Omega_1 \tag{2.21}$$

Here γ_a is the gyromagnetic ratio of the detected nucleus. Equation 2.21 states that the time/gradient strength at which the echo forms reflects the chemical shift Ω_1 of the initially encoded nucleus (figure 2.9 (3)). The detected time-domain signal therefore directly reflects the resonance frequencies along the spatially-encoded dimension without any Fourier transformation. After echo formation, the magnetization will be dephased over the sample volume again.

To monitor also the frequency of the detected nucleus, Ω_2, the gradient-assisted re-phasing and de-phasing can be repeated numerous times by oscillating the sign of the acquisition gradient (figure 2.10). During each such gradient an echo is detected, and the echoes are modulated by Ω_2. Fourier transformation of the data points acquired as a function of t_2 will thus reveal the full 2D spectrum in one single signal detection.

The reliance on identical spatial dependence of encoding and acquisition gradients is also the key for acquisition of higher-dimensional single-scan spectra. The spatial encoding, that was exemplified above for a z-gradient, can in principle be performed by a gradient of arbitrary geometry, and helical winding of spin packets can be done simultaneously along *independent* gradients. Using e.g. two different gradients with different geometries, two different frequency dimensions can be spatially encoded. The readout of such a double-encoded magnetization can then be performed by nested oscillation of the two gradients, i.e. one gradient is oscil-

lated multiple times within each oscillation period of the other gradient. In this way, 3D and higher-dimensional spectra can be recorded in a single scan (Shrot and Frydman (2003)). The drawback of this approach is that the nested oscillation of multiple gradients requires fast-switching gradients of different geometry. To circumvent these strong hardware demands, multiple frequency axes can be encoded with the *same* gradient. The resulting spectrum then contains multiple frequencies in a conflated frequency axis, or in other words, a projection of several dimensions onto one (compare section 2.2.4). This is an example of how different approaches for fast NMR can be combined together, a feature that will be found in later chapters of this thesis.

The advantage of single scan NMR is clearly its unique capability to obtain in principle any multidimensional spectrum in only one scan. In principle, it is thus the only technique allowing to study e.g. very fast (subsecond) kinetics, and efforts to this end have been reported (Gal et al. (2006); Shapira et al. (2004a)).

Among its current difficulties are sensitivity limitations: Of course, if the intrinsic sensitivity of an experiment under given conditions is insufficient in a single scan even in 1D, then a single-scan nD technique must fail.

But in addition, there are also sensitivity problems linked to the ultrafast acquisition scheme: Spatial encoding is prone to signal losses due to translational diffusion, and losses due to pulse imperfections during the encoding pulses.

An additional sensitivity issue comes from the data sampling scheme: In the presence of the acquisition gradient, the range of frequencies of spins located in different parts of the sample is very wide. Therefore, in order to detect signal from the whole sample volume, the receiver filter bandwidth, fbw, has to be opened to:

$$fbw = \gamma_a G_a L/2 \tag{2.22}$$

where L is the length of the detection volume. The noise is proportional to this band width, which often has to be set to several tens of kilohertz, and the signal-to-noise is therefore intrinsically lower in ultrafast experiments.

Combination with hyper-polarization techniques seems to be a very promising way to circumvent sensitivity limitations, as shown by recent development in this direction (Frydman and Blazina (2007); Shapira et al. (2004c)).

For practical applications, setting up ultrafast experiments is currently not routine, and the time spent for setting up the experiment is often orders of magnitude longer than the actual experiment.

2.2.7 Summary

Many different methods have been proposed to circumvent the sampling problem in NMR. The range of ideas is very large: from computational approaches in concert with non-linear data acquisition, to approaches that borrow concepts from fields as diverse as magnetic resonance imaging (ultrafast NMR) and astronomy (projection reconstruction).

Interestingly, many of these ideas are compatible with each other: Different encoding/processing schemes may be combined into one experiment, giving the experimentalist the freedom to take advantage of the strengths of different methods in terms of reduced experimental time and enhancement of sensitivity and resolution.

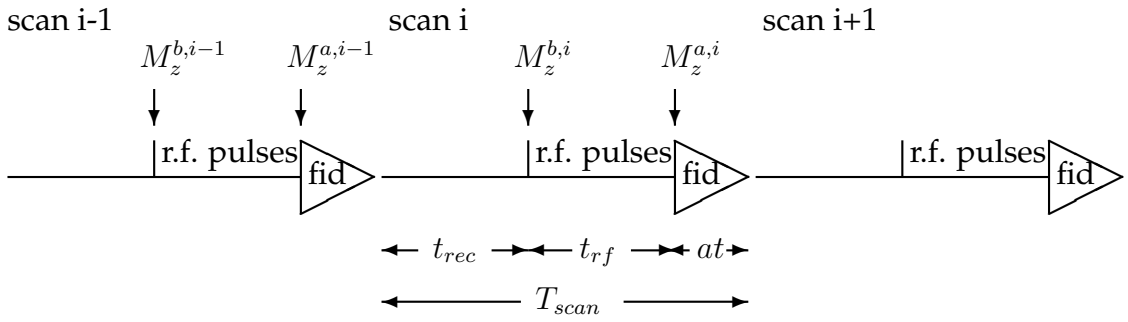
2.3 Accelerated nD NMR II: Fast pulsing techniques

A different way to acclerate the point-by-point sampling of the multidimensional time-space is to reduce the duration of the acquisition of each individual point.

This duration, T_{scan}, consists of the length of the pulse sequence t_{rf}, the acquisition period at and the delay between subsequent scans t_{rec}. The pulse sequence duration t_{rf} is determined by the information that is sought, and in the context of protein NMR experiments it is typically on the order of 100 ms or less. The duration of data acquisition, at, is determined by the desired resolution in the direct time domain and it was shown (equation 2.14) that is is generally limited by the T_2 time of the observed nucleus (generally $\leq$ 100 ms in proteins). These two durations are relatively short, and changing these periods would have severe consequences on the outcome of the experiment, which leaves the recovery delay, t_{rec} between two scans as the only parameter to be adjusted.

As mentioned in the discussion on page 38, the sensitivity is proportional to the magnetization present at the beginning of a given pulse sequence. The delay between consecutive scans allows this magnetization to be (partially) restored, and reducing the recovery delay will strongly influence the sensitivity. Before considering experiments that use reduced t_{rec} we should take a closer look how the sensitivity depends on the recovery delay. Again, the classical vector description of NMR suffices for this discussion.

Consider the following schematic pulse sequence:



The longitudinal magetization *before* a given scan i-1 is $M_z^{b,i-1}$. In the simplest (and most general) case, and neglecting losses during the pulse sequence, one can imagine this magnetization being rotated by a rf pulse, leaving after the pulse sequence a longitudinal magnetization $M_z^{a,i-1}$. The flip angle β of this rotation is given as $\cos\beta = M_z^b/M_z^a$. At the end of the pulse sequence, the longitudinal component is thus proportional to $\cos\beta$, whereas the transverse component is proportional to $\sin\beta$.

During the acquisition period and recovery delay the longitudinal magnetization is (partially) restored by relaxation processes leaving at the beginning of the subsequent scan i a magnetization $M_z^{b,i}$ along the z axis. Although the rate at which this relaxation occurs can be non-monoexponential, assume, for the sake of simplicity, a monoexponential recovery with a characteristic time constant T_1.

The longitudinal component present at the beginning of the subsequent scan, $M_z^{b,i}$, can then be expressed as:

$$M_z^{b,i} = M_0 - (M_0 - M_z^{a,i-1}) \cdot \exp(-\frac{at + t_{rec}}{T_1}) \tag{2.23}$$

where M_0 is the thermal equilibrium magnetization. For brevity, the period during which spins are allowed to relax freely, $at + t_{rec}$, is substituted in the following by T_{rec}. Under steady state conditions ($M_z^{b,i} = M_z^{b,i-1} = M_z^{b,eq}$) and keeping in mind that $M_z^a = M_z^b \cdot \cos\beta$, the magnetization available at the beginning of each scan $M_z^{b,eq}$ is thus given by:

$$M_z^{b,eq} = M_0 \, \frac{1 - \exp(-T_{rec}/T_1)}{1 - \exp(-T_{rec}/T_1)\cos\beta} \tag{2.24}$$

The detected signal, S, is proportional to the transverse magnetization at the end of the pulse sequence:

$$S \propto M_z^{b,eq} \cdot \sin\beta \tag{2.25}$$

Considering sensitivity, one must take into account that the noise adds up with the number of scans as $\sqrt{n}$. The number of scans that can be performed in a given

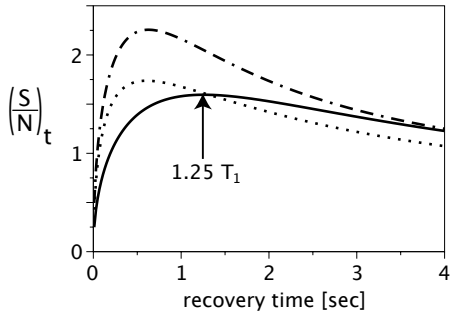


Figure 2.11 Simulated dependence of the sensitivity on the recovery delay T_{rec} and the amount of excitation, expressed as flip angle β, according to equation 2.27. The solid and dot-dashed lines assume $\beta = 90°$ and the dotted line was generated setting $\beta = 60°$. The relaxation time T_1 was set to 1 second (solid and dotted lines) and 0.5 seconds (dot-dashed line). The effect of the flip angle can thus be appreciated by comparing the solid and the dotted lines, and comparison between the solid and the dot-dashed lines demonstrates the sensitivity increase obtained from shorter T_1 relaxation. The maximum, found at $T_{rec} \approx 1.25 \cdot T_1$ when $\beta = 90°$ is indicated for the solid curve.

time (i.e. the repetition rate) depends on the inverse of the duration of one scan. The sensitivity is therefore proportional to

$$\left(\frac{S}{N}\right)_t \propto \sqrt{n} \propto \frac{1}{\sqrt{T_{scan}}} \tag{2.26}$$

Taking this equation together with equations 2.24 and 2.25 the sensitivity of a multi-scan experiment can be expressed as a function of the pulse sequence timing and the excitation angle as:

$$\left(\frac{S}{N}\right)_t \propto \frac{(1 - \exp(-T_{rec}/T_1)) \sin\beta}{1 - \exp(-T_{rec}/T_1)\cos\beta} \frac{1}{\sqrt{T_{scan}}} \tag{2.27}$$

Equation 2.27 is the key equation to keep in mind when trying to speed up data acquisition by reducing the recovery delay. It predicts the existence of an optimal recovery delay that maximizes sensitivity; for too short recovery delays spin relaxation is insufficient and for too long recovery periods too much time is wasted that could be used for accumulation of scans (see figure 2.11). Equation 2.27 reveals that for an excitation angle of $\beta = \pi/2$, optimal sensitivity is achieved setting

$$T_{rec}((S/N)_t^{max}) \approx 1.25 \cdot T_1 \tag{2.28}$$

Requirements for fast-pulsing experiments

The above considerations have shown how the sensitivity of an experiment depends on the interscan delay t_{rec}. A simple reduction of this delay comes at a price

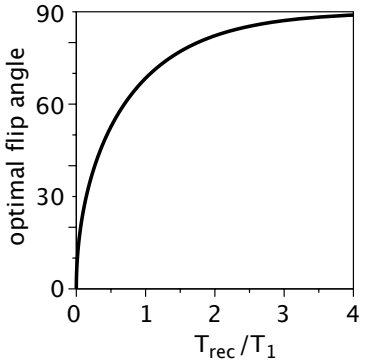


Figure 2.12 Optimized excitation angle as a function of the relative interscan delay, T_{rec}/T_1, calculated according to equation 2.29.

in terms of sensitivity, and it was found that for very short interscan delays the sensitivity curve drops steeply (figure 2.11). When studying biomolecules, concentration is limited and transverse relaxation losses are prominent, such that sensitivity is generally precious. The drop in sensitivity upon reducing the interscan delay is therefore most often not acceptable.

Methods that aim at accelerating data acquisition by reducing the recovery delay necessarily have to increase the steady-state longitudinal magnetization in the fast-pulsing regime. In this context, different methods can be envisaged and the remainder of section 2.3 will discuss these different strategies.

2.3.1 Partial excitation and the Ernst angle

Ernst and co-workers have developed an elegant way to optimize sensitivity in simple *pulse - acquire* one-dimensional experiments when applied at rapid repetition rates. Applying the excitation pulse with an optimized effective excitation angle (β_{Ernst}) accounts for the incomplete recovery of magnetization. It can be shown (Ernst and Anderson (1966); Ernst et al. (1987)) that optimal sensitivity is achieved when:

$$\cos \beta_{Ernst} = \exp(-T_{rec}/T_1) \tag{2.29}$$

Figure 2.12 shows the optimal flip angle (Ernst angle) as a function of the recovery delay. Ninety degree excitation is the method of choice only using a recovery delay that is at least 4 times longer than T_1, whereas smaller effective flip angles are best used to account for incomplete magnetization recovery in the fast-pulsing regime.

The gain in sensitivity that is obtained when using a non-90° degree excitation

for fast repetition rates is shown in figure 2.11 (solid versus dotted line).

In conclusion, an optimization of the flip angle yields significant advantages in sensitivity when fast repetition of the experiment is desired. However, in complex pulse sequences for multidimensional NMR, implementation of non-90° degree excitation is generally not straightforward. Some possible methods are shown in chapter 3 (section 3.1.1).

2.3.2 Accelerating the magnetization recovery by optimized pulse sequence design

Increasing the repetition rate of an experiment is in general accompanied by a decrease in sensitivity. In principle, every (spectroscopic) method that increases sensitivity can therefore be of interest in the context of fast-pulsing techniques. The range of such spectroscopic methods is very wide, from planar-mixing sensitivity enhanced methods (Cavanagh et al. (1991); Cavanagh and Rance (1990); Kay et al. (1992); Palmer et al. (1991)) to transverse relaxation optimized methods (Fiala et al. (2000); Pervushin et al. (1997); Sklenar et al. (1993)) and many other pulse sequences that optimize coherence transfer or relaxation properties. These methods generally yield an overall increase of the sensitivity that is independent of the repetition rate. Discussing all such methods is not within the scope of the current thesis and the current discussion will only focus on concepts that have been introduced in the context of fast-pulsing techniques.

As can be appreciated from equation 2.27, increasing the speed of recovery of magnetization (loosely referred to as time constant T_1 in equation 2.27) increases the sensitivity in the fast-pulsing regime. This is shown in figure 2.11 (compare solid and dot-dashed lines).

2.3.2.1 Strategies

Two groups of strategies to achieve fast magnetization recovery can be distinguished:

On the one hand, adding relaxation agents to the sample can be shown to yield significant sensitivity gains under certain conditions. Paramagnetic relaxation agents make use of the large dipole moment of unpaired electrons to enhance nuclear spin longitudinal relaxation. Different approaches have been introduced, either aiming at directly accelerating the relaxation rate of the observed nuclei (Cai et al. (2006); Eletsky et al. (2003)), or by enhancing the rate of water

^{1}H longitudinal relaxation (Hiller et al. (2005b)), which is of advantage in the case of exchange of protein-bound protons with water protons. These methods are not within the scope of this work, and theory and applications can be found in the literature pointed to above, as well as in a recent review on paramagnetic relaxation agents (Caravan et al. (1999)).

On the other hand, there are spectroscopic methods that aim at achieving a rapid magnetization recovery by optimized pulse sequence design. In the context of biomolecules, longitudinal relaxation enhancement of ^{1}H spins, developed in the present thesis, has proven very successful. This section (2.3.2) will discuss the principles underlying this method, especially for application to protein NMR spectroscopy.

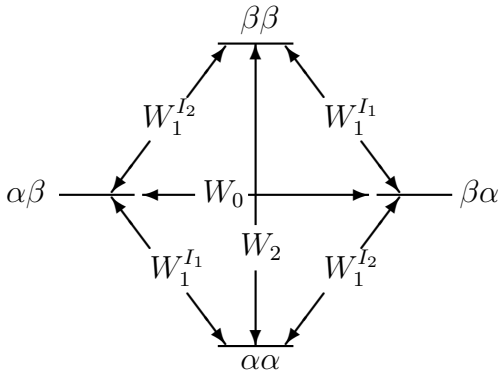
2.3.2.2 Proton relaxation enhancement by dipolar cross-relaxation: Principles

Virtually all the standard experiments performed in liquid state NMR to study diamagnetic protein molecules excite and detect proton spins due to their elevated gyromagnetic ratio and the corresponding sensitivity advantages (see equation 2.5 on page 34). In a multiscan experiment, it is therefore the recovery of proton magnetization that is important for the sensitivity.

Proton longitudinal relaxation in diamagnetic proteins is dominated by the dipole-dipole interaction between ^{1}H spins. A detailed description of relaxation theory can be found in many textbooks (see e.g. Cavanagh et al. (1995)). Only some relevant aspects of the theory of dipole-dipole relaxation are introduced in the next sections, and the consequences for the relaxation in proteins will then be studied by simulation and experiment.

Proton longitudinal relaxation basics: the two-spin system

To introduce some of the concepts that govern dipolar relaxation, it is useful to re-examine a simple system containing only two dipolar coupled spins $\frac{1}{2}$, termed I_1 and I_2. Each of these two spins has two eigenstates, termed α and β. Four eigenstates of the coupled spin system exist therefore, and these are donted $\alpha\alpha$, $\beta\alpha$, $\alpha\beta$ and $\beta\beta$. These four eigenstates in general have different energies, with the central two states, $\alpha\beta$ and $\beta\alpha$, being near-degenerate for a homonuclear system at high magnetic field. For nuclei with positive γ the lowest-energy state is the state $\alpha\alpha$. The following scheme summarizes this situation.



Each of these four energy eigenstates has a population associated with it (expressed as n_{lk} for the state lk), and under equilibrium conditions it is given by the Boltzmann distribution (equation 2.2). The longitudinal magnetization of the two spins, I_{1z} and I_{2z} is given by the differences between the populations of the α and β states for each of the spins as follows:

$$\begin{aligned} I_{1z} &= \frac{1}{2}\, n_{\alpha\alpha} - \frac{1}{2}\, n_{\beta\alpha} + \frac{1}{2}\, n_{\alpha\beta} - \frac{1}{2}\, n_{\beta\beta} \\ I_{2z} &= \frac{1}{2}\, n_{\alpha\alpha} + \frac{1}{2}\, n_{\beta\alpha} - \frac{1}{2}\, n_{\alpha\beta} - \frac{1}{2}\, n_{\beta\beta} \end{aligned} \tag{2.30}$$

Transitions between the four eigenstates involve the change in the state of either one spin or both spins. The corresponding transition probabilities are expressed as W. Transitions involving only one spin are referred to as single-quantum transitions (and the corresponding transition probability is given as W_1), whereas transitions involving two spins can be either zero-quantum ('flip-flop') transitions, with the transition probability W_0, or double quantum ('flip-flip') transitions (W_2). The situation is summarized in the above energy-level diagram.

The energy associated with the single-quantum transitions of spin I is given by the energy difference between the two involved states, which corresponds to the Larmor frequency of this spin. Zero- and double quantum transitions involve transition energies that correspond to the difference and the sum of the Larmor frequencies.

Transitions between the eigenstates involve energy exchange with the environment. The spin system thus has to be coupled by some way to the surrounding, and it is this surrounding that either takes up energy or makes energy available to the spin system.

It can be shown that radiation-less processes constitute the mechanism for this

energy exchange, and the energy transfer in these processes occurs between the spin system and the motional degrees of freedom (Abragam (1961)). To fix ideas: The two-spin system that we are looking at is embedded in a molecule that has some motional degrees of freedom given by overall and internal dynamics (stochastic motions). Because of these dynamics, each of the two spins 'sees' an effective magnetic field (arising from the other spin) that fluctuates with time, and it is this fluctuation that causes spin relaxation (dipole-dipole relaxation). The dynamics are thus induced by the molecular motion, which constitutes the energy bath that the spin transitions are coupled to. This energy bath is called 'lattice'.

As stated above, the spin transitions require a certain energy from the lattice that matches the transition energy. In other words, the stochastic fluctuation of the magnetic field experienced by a spin has to occur at the 'right' frequency. The efficiency of the spin transitions is correlated to the amount of energy that is present in the lattice at the relevant frequencies. The amount of motional energy as a function of the frequency of the motion is quantified by the spectral density function $J(\omega)$. A detailed introduction and definition of the spectral density can be found in many textbooks (e.g. Abragam (1961); Cavanagh et al. (1995)), and is not within the scope of this PhD thesis. Here, it should suffice to note that the spectral density reflects the amount of random fluctuation as a function of the frequency of this random fluctuation. It can be modeled making assumptions about the motions present in the molecule, as shown below for the case of isotropic reorientational motion.

We have now the basic tools and concepts at hand to come back to the relaxation in the two-spin system. As seen from equation 2.30, the longitudinal magnetization of the two spins can be expressed as differences in the populations of the energy levels. The time evolution of the population of a given energy level can be deduced from the transition probabilities *to this given energy level* from all the other levels and the transition probabilities *from this level* to all other energy levels, as well as the populations of all other energy levels. For example, the evolution of the population of the level $n_{\alpha\alpha}$ can be written as[4]:

$$\frac{\mathrm{d}}{\mathrm{d}t} n_{\alpha\alpha} = -W_1^{I_1}\, n_{\alpha\alpha} - W_1^{I_2}\, n_{\alpha\alpha} - W_2\, n_{\alpha\alpha} + W_1^{I_2}\, n_{\beta\alpha} + W_1^{I_1}\, n_{\alpha\beta} + W_2\, n_{\beta\beta} \qquad (2.31)$$

Similar equations can be found for all possible energy eigenstates ($n_{\alpha\beta}$, $n_{\beta\alpha}$, $n_{\beta\beta}$). Taking all these relations together with equation 2.30 the longitudinal relaxation of

[4]To be more exact, thermally corrected transition probabilities should be used here, that account for the fact that in thermal equilibrium the energy levels have different populations.

the two spin system can be expressed by the following equations:

$$-\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} I_{1z}-I_{1z}^0 \\ I_{2z}-I_{2z}^0 \end{pmatrix} = \begin{pmatrix} \rho & \sigma \\ \sigma & \rho \end{pmatrix}\begin{pmatrix} I_{1z}-I_{1z}^0 \\ I_{2z}-I_{2z}^0 \end{pmatrix} \tag{2.32}$$

Here I_{1z}^0 is the thermal equilibrium longitudinal magnetization of spin I_1, ρ is the auto-relaxation or leakage rate constant and σ is the cross-relaxation rate constant. The above relations (2.32) are known as the *Solomon equations* for a two-spin system.

They predict that the recovery of the longitudinal spin magnetization on one spin (I_1) depends on this spin's deviation from thermal equilibrium ($I_{1z}-I_{1z}^0$) with a rate constant of ρ, but also on the deviation of the *other spin* from its thermal equilibrium magnetization ($I_{2z}-I_{2z}^0$), involving the rate constant σ. The former process is termed *auto-relaxation*, whereas the latter is designated as *cross relaxation*. Therefore, even in this very simple case of only two interacting spins, the relaxation of one spin cannot be reflected by a single longitudinal relaxation time constant.

The rate constants, ρ and σ, depend on the transition probabilities W_0, W_1 and W_2. These in turn depend on the distance r between the spins, their gyromagnetic ratios (assumed here to be the same, homonuclear case) and the spectral densities at the frequencies 0 (the difference frequency), ω (the Larmor frequency which is common to both spins) and 2ω (the sum frequency). The rate constants are given by[5]:

$$\begin{aligned} \rho &= W_0 + 2W_1 + W_2 = \frac{1}{10}\, d^2\, (J(0) + 3J(\omega) + 6J(2\omega)) \\ \sigma &= W_2 - W_0 = \frac{1}{10}\, d^2\, (6J(2\omega) - J(0)) \qquad (2.33) \\ \text{with} \quad d &= \frac{\mu_0}{4\pi}\frac{\hbar\gamma^2}{r^3} \qquad (2.34) \end{aligned}$$

where γ is the gyromagnetic ratio of ^{1}H and μ_0 is the magnetic permeability of free space. The spectral density function describes the dynamics of the system, which is in general not known in detail. However, it can be modeled either assuming a certain form of the motion or extracting information from molecular dynamics simulations, or simply assuming a functional form of the spectral density function (without specifying the motional model causing the given functional form). For example, one often assumes a very simple model, where isotropic tumbling of the

[5]Note that in the literature σ is sometimes found replaced by -R_{cross} thus with inverted sign.

molecule is the only motion that modulates the orientation (but not the distance) between the two spins. The spectral density function in this case is then given by

$$J(\omega) = \frac{\tau_c}{1 + \omega^2 \, \tau_c^2} \tag{2.35}$$

The resulting rate constants ρ and σ are shown as a function of the characteristic time constant of this random reorientational motion (the correlation time τ_c) in figure 2.13 (a). Whereas the auto-relaxation rate constant is positive irrespective of the correlation time, the cross-relaxation rate σ is positive only in the case of fast reorientational motion and becomes negative for larger τ_c. As can easily be derived from equations 2.35 and 2.33 the zero-crossing is found when $\omega\tau_c = \sqrt{5}/2 \approx 1.12$. Proteins under generally used conditions (solvent, temperature, magnetic field strength) basically always have τ_c values above this zero-crossing point and cross-relaxation rate constants are thus generally negative. As will be shown below, 'flip-flop' processes become very efficient in this regime and this gives rise to a process called spin diffusion (see below).

It is instructive to examine the relaxation of the two-spin system in two cases of interest.

Relaxation after non-selective inversion of both spins. The initial conditions in this situation are $I_{1z} = -I_z^0$ and $I_{2z} = -I_z^0$ and the solution of the Solomon equations is:

$$\begin{aligned} I_{1z} - I_z^0 &= -2 \, I_z^0 \, \mathrm{e}^{-(\rho+\sigma)t} \\ I_{2z} - I_z^0 &= -2 \, I_z^0 \, \mathrm{e}^{-(\rho+\sigma)t} \end{aligned} \tag{2.36}$$

Relaxation after selective inversion of spin I_2. The initial conditions in this situation are $I_{1z} = I_z^0$ and $I_{2z} = -I_z^0$ and the longitudinal magnetizations of the two spins will evolve according to:

$$\begin{aligned} I_{1z} - I_z^0 &= -I_z^0 \, (\mathrm{e}^{-(\rho+\sigma)t} - \mathrm{e}^{-(\rho-\sigma)t}) \\ I_{2z} - I_z^0 &= -I_z^0 \, (\mathrm{e}^{-(\rho+\sigma)t} + \mathrm{e}^{-(\rho-\sigma)t}) \end{aligned} \tag{2.37}$$

The behavior of the magnetization for these two cases is illustrated in figure 2.13 for positive σ (c, d) and negative σ (e, f).

Whereas in the case of non-selective inversion the relaxation of both spins is given by a single time constant ($\rho + \sigma$), in the case of selective I_2 inversion an additional exponential component in I_2's recovery appears that evolves with a rate ($\rho - \sigma$). Whether this component is faster or slower than the one given by $\rho + \sigma$

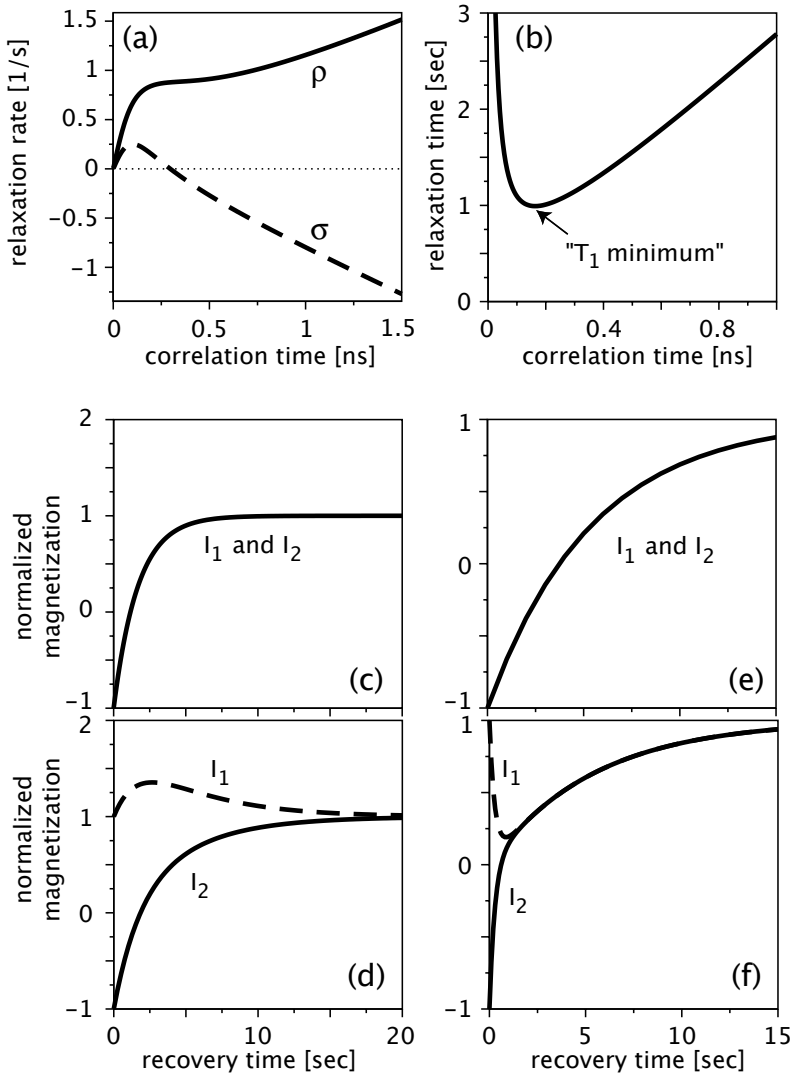


Figure 2.13 Dipolar relaxation in a two-spin system undergoing isotropic tumbling. (a) Relaxation rate constants ρ (solid line) and σ (broken line) as a function of the isotropic tumbling correlation time τ_c for a pair of proton spins separated by 0.2 nm in a magnetic field of 14.1 T (600 MHz ^{1}H Larmor frequency). (b) Relaxation time constant $T_1 = 1/(\rho+\sigma)$ in the case of non-selective inversion of the two spins (see equation 2.36). A minimum T_1 is found for $\tau_c \approx 0.17$ ns (conditions as in (a)). (c, d) Time course of the magnetizations in this two-spin system at 14.1 T following non-selective inversion (c) and selective inversion (d) of spin I_2. The correlation time is set to 50 ps (positive σ regime). Note that the recovery of spin I_2 is slower in the selective case than in the non-selective one. (e, f) As (c, d) but τ_c is set to 2 ns (negative σ regime). Note the faster effective recovery of spin I_2 in the selective case compared to non-selective inversion.

depends on the sign of σ (equation 2.33). When the dynamics is fast (small molecule, low viscosity) then selective inversion of one spin leads to a transient hyperpolarization of the other spin; this causes a retarded relaxation of the inverted spin (figure 2.13 (d) with respect to the situation of non-selective inversion.

In contrast, for a slowly-tumbling spin pair (e.g. in a macromolecule) $\sigma < 0$ and in this case the additional component found in the selective case is faster. It causes a rapid (partial) recovery of the inverted spin which decreases the magnetization of the initially unperturbed spin (figure 2.13 (f). Both spins then fully recover in one single phase.

Let us briefly summarize the conclusions from this section:

- Even in the very simple case of a dipolar coupled two-spin system the relaxation upon excitation can in general not be expressed by a single time constant. The relaxation of each spin depends on the state of the other spin (equation 2.32).
- The sign of the cross relaxation constant σ is determined by the spectral density function at frequencies 0 and $2\omega_H$ (equation 2.33). Assuming that relaxation is mediated by an isotropic reorientational motion characterized by a single time constant τ, the cross-relaxation rate constant is negative for values of $\omega\tau$ above $\sqrt{5}/2$ (≈ 1.12), which is generally the case for macromolecules at high fields (*negative σ regime*).
- In the negative-σ regime, a fast-recovery component is found when a spin is selectively inverted (while leaving the other spin in equilibrium).

The realization that in macromolecules selectively excited/inverted spins have a recovery component with a higher rate, as compared to the non-selective case, is at the heart of what has become known as *longitudinal relaxation optimized* methods in biomolecules.

Proton longitudinal relaxation in proteins: Theory

Protein molecules contain a large number of proton spins. The dipolar relaxation between them underlies the same principles as in the two-spin system, and the time evolution of the longitudinal magnetization of each spin in a molecule depends on the instantaneous magnetization of all other proton spins. The time evolution of the magnetization of the n ^{1}H spins (I_1 to I_n) is given by the following differential equations (Solomon equations):

$$-\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} I_{1z} - I_{1z}^0 \\ I_{2z} - I_{2z}^0 \\ \vdots \\ I_{nz} - I_{nz}^0 \end{pmatrix} = \begin{pmatrix} \sum_j \rho_{1j} & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & \sum_j \rho_{2j} & \dots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \dots & \dots & \sum_j \rho_{nj} \end{pmatrix} \begin{pmatrix} I_{1z} - I_{1z}^0 \\ I_{2z} - I_{2z}^0 \\ \vdots \\ I_{nz} - I_{nz}^0 \end{pmatrix} \tag{2.38}$$

In addition to the ^{1}H spins bound to the molecule, in aqueous solution another important group of proton spins has to be considered: the proton spins of the solvent water molecules. Water proton spins can interact with protein protons and thus influence the relaxation of the latter in two ways: They can have a direct cross-relaxation (nOe) through space, or they can act via a more indirect mechanism, namely if a proton bound to the protein is replaced by a water proton spin (water exchange of labile hydrogen sites).

The first of these two mechanisms is relatively inefficient and the direct water-protein ^{1}H-^{1}H cross relaxation rates are small. This is explained by the fact that the residence time of solvation water molecules around proteins is very short, typically on the subnanosecond time scale (Otting et al. (1991); Wuthrich et al. (1992)).

The second mechanism, exchange of labile hydrogen atoms in the protein for water protons can cause more pronounced effects, as investigated by simulation below.

Proton longitudinal relaxation in proteins studied by simulation and experiment

What we initially attempted was to find a way of accelerating the ^{1}H magnetization recovery. The next sections investigate conditions where this is the case.

A first idea of how the magnetization recovery can be accelerated comes from the above analysis of the two-spin system. Figure 2.13 (f) showed that in the two-spin system the recovery of one spin can be significantly faster when this spin is *selectively* excited, in comparison to excitation of both spins. This could potentially be of interest for enhancing longitudinal relaxation and thus sensitivity in fast-pulsing experiments for proteins.

However, before going deeper into this subject we should briefly stop and analyze if selective-excitation experiments could in principle be feasible and of any use for applications to proteins.

Selectively exciting part of the proton spins while leaving part of them unperturbed of course has a price: Those spins that are not excited will ultimately not give any signal. So are there any situations where one can afford *not* exciting all the proton spins in the molecule of interest ?

The study of a protein by NMR requires collecting a large amount of information: one wants to assign the resonance frequencies of - if possible - all the present nuclei, collect distance and angle information, measure coupling constants and relaxation rates. The perfect experiment would offer all this information at a time. In practice, however, this is not feasible: obtaining a specific piece of information requires an experiment that has an accurately timed train of pulses and delays optimized such that the magnetization follows the desired pathway and ultimately gives the information one is looking for. It is therefore impossible to collect all different kinds of information with the same experiment. In practice one records different experiments: experiments that allow assigning the resonances of the backbone atoms, those that allow connecting the backbone atoms to the sidechains, homonuclear experiments, heteronuclear ^{1}H-^{13}C or ^{1}H-^{15}N correlation experiments etc. Among all these experiments there are (many) experiments that exploit only a subset of the proton spins: the mentioned ^{1}H-^{13}C or ^{1}H-^{15}N correlation experiments exploit only the magnetization of proton spins having a scalar coupling to ^{13}C or ^{15}N, respectively; experiments that focus on aromatic side chains only utilize magnetization residing on aromatic protons, many backbone experiments exploit only amide proton magnetization.

Taken together, there are many instances where a selective excitation of only part of the proton spins suffices to get the information sought. It is therefore potentially of practical relevance to study in more detail proton relaxation under conditions of selective or non-selective excitation and see what can be learned about it and how one can benefit from selective schemes in cases where rapid relaxation is sought.

The next sections focus on the relaxation of backbone amide protons, as this is the group of protons that is exploited for backbone assignment experiments, an important step in every atom-resolved study of protein structure and dynamics. In addition, the amide protons are attractive for selective excitation schemes because their resonance frequenies are generally well-separated from the rest of the proton spins making selective manipulation by shaped rf pulses feasible. The relaxation of these protons will be studied by simulation and experiment, which will then reveal strategies for fast-pulsing, sensitive experiments.

Amide ^{1}H relaxation in proteins studied by simulation and experiment

A simulation program was written as described in the appendix (section 10.1 on page 215). Briefly, the algorithm retrieves structural coordinates from a pdb file and calculates the evolution of the magnetization of all proton spins present in the molecule using the following assumptions:

- The spectral density function used for the calculation of the relaxation rates assumes isotropic overall tumbling of the molecule with a correlation time τ_c. In addition, fast rotation of the methyl groups has been implemented to study the effect of internal dynamics on the ^{1}H relaxation.

- Exchange of labile sidechain atoms has been implemented and simulations can be performed either assuming that water ^{1}H magnetization is in thermal equilibrium or equal to zero and recovering its equilibrium value with a time constant of 3 seconds. The exchange of labile sidechain hydrogens is assumed to be very fast. Exchange of amide protons with water ^{1}H has been neglected. This is justified by the fact that in folded proteins at neutral (or lower) pH exchange of amide protons generally occurs at much longer time scales.

- Proton longitudinal relaxation generally has a non-monoexponential behavior. In order to express the potentially complex magnetization recovery of a given spin by a single value, an effective T_1 value is given as the time where the magnetization reaches $\sim$ 63% of its thermal equilibrium value. (This is based on the fact that for a monoexponential recovery the recovered fraction at the characteristic time is $1 - 1/e$=63%.)

To evaluate the simulation program, experimental and simulated inversion recovery curves were compared. Figure 2.14 shows the recovery of the longitudinal magnetization of some representative amide protons in the molecule ubiquitin (8.6 kDa) as revealed by experiment and simulation. Panels (a) and (b) correspond to non-selective and amide ^{1}H selective inversion, respectively. For the case of selective amide spin inversion, figure 2.14 (c) shows the time course of the magnetization of some representative aliphatic proton spins. The data can be summarized as follows:

Amide spin relaxation in a compact, folded protein has a different behavior in the cases of selective and non-selective inversion. The recovery after non-selective inversion follows to a good approximation a monoexponential behavior, whereas two exponential components are apparent in the recovery after selective amide spin inversion. Due to the additional fast component almost-complete restoration of magnetization is achieved in much shorter time. The fast component in the

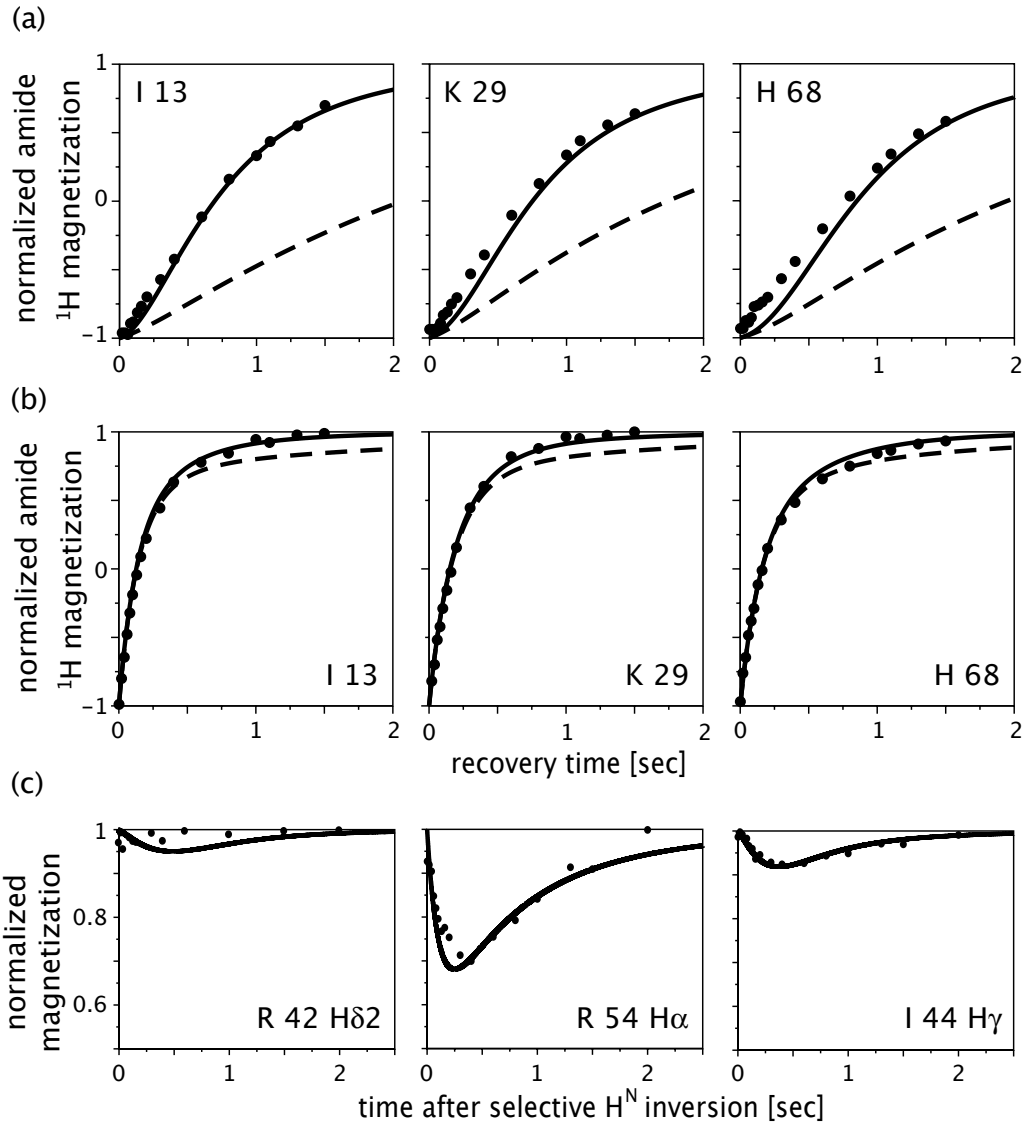


Figure 2.14 Time evolution of the longitudinal magnetization of some representative ^{1}H spins revealed by simulation and experiment. (a) Inversion-recovery of amide protons of residues 13, 29 and 68 after non-selective inversion. Filled circles: experiment, solid line: simulation with methyl rotation, dashed line: simulation without methyl rotation. (b) Recovery after selective amide ^{1}H inversion. (c) Evolution of the magnetization of three representative aliphatic ^{1}H spins after selective amide ^{1}H inversion. Again, circles and solid lines correspond to experimental and simulated (with methyl rotation) data. In all plots, the normalized magnetization ranges from 1 (thermal equilibrium magnetization) to -1 (inversion of the thermal equilibrium populations). Details about the experiment and the simulation are given in section 10.2 on page 217.

recovery of amides upon selective excitation can be attributed to cross-relaxation to aliphatic protons. These aliphatic proton spins are driven away from their initial equilibrium state with a time constant that is on the order of the time constant of the fast amide relaxation component (panel (c)). The recovery of the aliphatic ^{1}H magnetization then follows the same kinetics as the slow component of the amide ^{1}H relaxation, in a similar way as was found for the two-spin system.

The time course of longitudinal amide spin magnetization can be predicted very well with the simple model used, where methyl rotation is the only mode of internal dynamics that is considered, as can be appreciated from the good agreement between experimental (circles) and simulated data (solid curves). However, suppression of the methyl dynamics yields very unsatisfactory predictions (broken lines). This points to the methyl groups as 'relaxation sinks' in proteins, a fact that has been predicted and demonstrated experimentally before (Akasaka et al. (1990); Olejniczak and Weiss (1990)). Despite the good general agreement, especially found for amide protons, the simulations fail to correctly reproduce some experimental recovery curves found for aliphatic ^{1}H spins and ^{1}H spins located in dynamic regions (see below).

It is interesting to note also that ignoring the methyl rotation has a much more drastic effect on the amide relaxation in the non-selective case than in the selective case. In the selective case the simulated recovery upon suppression of the CH_3 rotation is still found bi-exponential and the fast component is reproduced well. In contrast, the predicted slow component is slower than found experimentally. In the non-selective case ignoring the methyl relaxation yields amide relaxation rates that are much too slow.

How can we interpret these findings ? The slow overall tumbling of the molecule (macromolecular case) makes cross-relaxation ('flip-flop' or 'spin diffusion') processes very efficient: upon selective excitation the magnetization is quickly spread over the whole proton spin system. *Evacuation* of the excitation energy from the spin system (rather than redistribution), is mediated by the fast dynamics of the methyl groups. In the case of selective excitation, the fast component corresponds to the rapid redistribution of excitation energy whereas the second component can be attributed to the evacuation of the energy from the spin system (compare this behavior to the one found for the two-spin system, figure 2.13(d)). Suppressing the methyl rotation therefore leads to a reduction of the rate of this second component, while the fast component is reproduced well.

In the case of non-selective inversion of all proton spins, the situation is different: flip-flop processes are effectively quenched, because all proton spins are initially in an excited (inverted) state. The rate of recovery of the amide proton spins is therefore given by the evacuation of the energy from the whole spin system, which is mediated by fast dynamics (in a similar way as the slow component

found upon selective inversion). Suppression of the fast methyl rotation in the non-selective case therefore leads to dramatically increased relaxation times.

Amide relaxation and water magnetization

Some of the oxygen- or nitrogen-bound sidechain protons in proteins as well as backbone amide protons can exchange with water, i.e. the bond to the O or N atom in the protein can be broken and a water proton takes the place instead. Such exchange of labile protons for water protons introduces spins into the protein that are in the magnetization state of the water protons. This is a way how the magnetization of ^{1}H from water interferes with the protein.

The longitudinal relaxation of water protons is generally longer than the one of protein-bound protons. If water-^{1}H spins as well as protein-^{1}H spins get excited by a given pulse sequence then the exchange of protein-bound spins for water-^{1}H introduces protons that deviate more from the equilibrium state than the protein-bound protons. This will in turn slow down longitudinal relaxation of the latter.

In contrast, if the water proton magnetization is close to Boltzmann equilibrium then the same process would introduce a "relaxed" proton spin which will then in turn cross-relax neighboring protein-bound protons. Spin diffusion then propagates the magnetization of this newly introduced proton all over the protein.

The higher the efficiency of spin diffusion, the more pronounced is the effect of such water exchange for the whole protein molecule. We will thus return to this issue below when considering the dependence of the ^{1}H relaxation on the protein size.

The concentration of water in aqueous protein samples is $\approx 10^4$ times larger than the protein concentration and suppression of the water signal is indispensable in order to observe protein signal. Use of selective pulses or tailored sequences of hard pulses (Piotto et al. (1992)) allows to selectively manipulate water magnetization making possible the suppression of water signal without affecting protein signal. A class of such suppression techniques makes use of magnetic field gradients to dephase the bulk water magnetization that is in an excited state. However, from the above consideration it is evident that such techniques, that result in an excitation of water, can lead to sensitivity losses. Experimental schemes that leave the water magnetization close to thermal equilibrium have therefore been developed and are widely implemented in many biomolecular NMR pulse sequences (Grzesiek and Bax (1993); Redfield et al. (1975); Stonehouse et al. (1994)).

However, no experimental flip-back scheme is perfect, and generally only a fraction of the water ^{1}H magnetization gets restored at the end of a scan. When

repetition rates faster than the water ^{1}H longitudinal relaxation rate are used, small imperfections can lead to significant deviations of the steady-state water magnetization from thermal equilibrium. Clearly, clean water flip-back schemes are required to maximize the steady-state water ^{1}H magnetization. In this context acceleration of the water recovery by use of paramagnetic molecules has been proposed (Hiller et al. (2005b)).

The simulations below consider the extreme cases of either leaving water ^{1}H magnetization in the thermal equilibrium state or exciting bulk water magnetization. Any real case is in between these two scenarios.

Dependence on the size of the molecule

The preceding simulations and experiments have revealed that for a relatively small protein like ubiquitin, dramatically accelerated amide ^{1}H magnetization recovery is found when only the amide spins are excited selectively, while leaving the aliphatic, aromatic and water protons spins in thermal equilibrium. The mechanism for this enhanced recovery is spin diffusion and equations 2.33 and 2.35 predict that the cross-relaxation rate giving rise to spin diffusion depends on the overall tumbling correlation time. In addition, the number of coupling partners available, and thus the number of relaxation sources, can also be expected to influence the effective relaxation rate.

One might therefore ask the question how the relaxation behaves for smaller or larger systems. Is there a minimal size necessary to observe significant enhancements in the selective scheme?

This issue can be divided into two separate questions. First, how does the number of spins present in the system (the number of possible acceptors or relaxation sources) influence the relaxation ? And second, what is the dependence of the effective relaxation rates on the overall dynamics, which is correlated to the molecular size and shape ?

To shed light on the first question, simulations were performed as follows: An amide proton buried deep inside a large molecule was studied (H^N of residue 271 in malate synthase G (Tugarinov et al. (2005)). It was taken as the center of spheres of different radii. The radius of these spheres was varied and only the proton spins located inside these spheres were taken into account for the simulations. The relaxation of the central amide proton was studied as a function of the radius of the spheres, while the overall correlation time was held constant (5 ns). Of course, a constant tumbling correlation time with varying size of the molecule is physically unrealistic. However, this simulation aims at separating the effect of the number

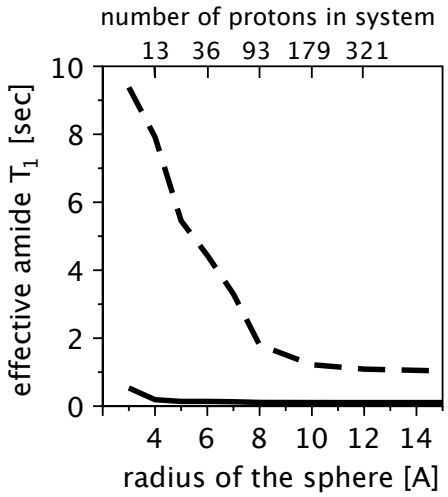


Figure 2.15 Effective relaxation rate of an amide proton in the center of spheres of tightly packed protein environment as a function of the radius of the spheres. The relaxation time constants in the case of selective amide excitation and hard excitation are indicated by solid and dashed lines, respectively.

of spins from the strength of the coupling between them, which is related to the tumbling correlation time.

Figure 2.15 shows the results of these simulations. For the case of selective excitation (solid curve), the effective relaxation time is very insensitive to the number of spins and a very small number of spins seems to be sufficient to insure fast recovery of the studied proton spin. This is expected from the behavior of the two-spin system (figure 2.13): distributing the excitation energy to even a relatively small number of spins greatly enhances the relaxation rate. In the non-selective case, more spins are needed to relax the spin under investigation.

In both cases, a plateau level is reached for protein fragments with a radius of more than 9 Å. Even very small proteins are well above this size, indicating that the number of protons present, even in small proteins, is not the limiting factor for proton relaxation.

The second factor related to the size of the molecule is the overall tumbling correlation time τ_c. This factor directly alters the coupling strength between spins (see equation 2.33 and 2.35 on page 70). Figure 2.16 shows the average effective relaxation times in the protein villin HP36 (McKnight et al. (1997), pdb entry 1VII, 4 kDa) as a function of τ_c for the case of selective excitation of amide protons only, for non-selective excitation of protein-bound ^{1}H spins with water flip-back and for non-selective excitation without water flip-back. The left panel shows the range of

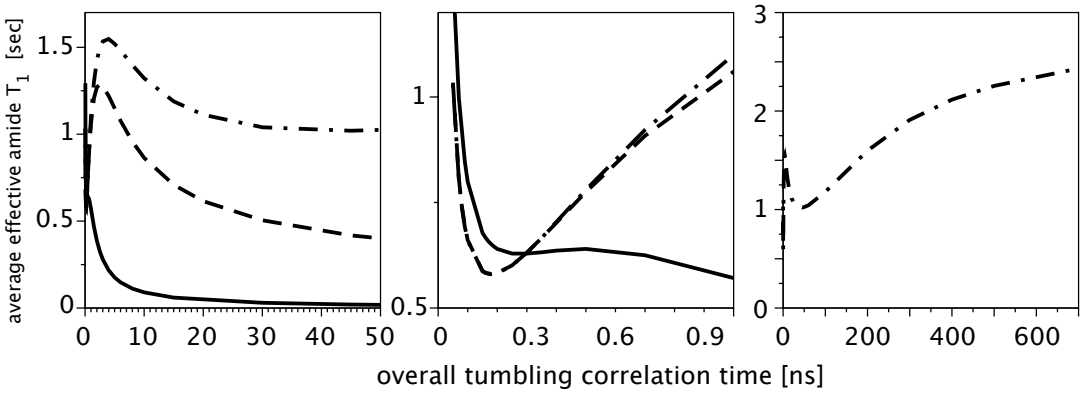


Figure 2.16 Average effective amide ^{1}H relaxation times for the protein villin-HP36 as a function of the correlation time for overall tumbling, τ_c. The simulations were done for 600 MHz ^{1}H frequency, methyl rotation was included as before. The average relaxation time for all amide protons is shown for the case of amide-selective (solid line), non-selective excitation with water flip-back (dashed line) and non-selective excitation without water flip-back (dash-dotted line). Water exchange was included as described above (page 216). The middle panel is a zoom in the low-τ_c region and the right panel shows the raise of the relaxation rates in the case of hard excitation without water flip-back in the limit of very high τ_c.

τ_c values typically encountered in biomolecular studies from a few nanoseconds up to some tens of ns. The central and right panel focus on the sub-nanosecond- and high-τ_c range, respectively.

These curves are corroborated by experimental data, shown in figure 2.17 for two proteins with tumbling correlation times of approximately 4 and 8 ns, respectively. The curves show the sensitivity as a function of the interscan delay. As shown by equation 2.28 and figure 2.11, the form of these curves depends on the relaxation time constant T_1. For 90° excitation (which is the case here) the maximum values are found at a recovery time corresponding to 1.25 times T_1. The larger SiR-FP18 shows shorter effective T_1 times for both the selective and the non selective water flip-back (wfb) experiment, in a manner that agrees qualitatively well with the simulations.

It is instructive to take a closer look at the behavior of these curves to get a somewhat deeper understanding of the relaxation behavior of coupled spin systems. It is also useful to remember the behavior found for the two-spin system (see figure 2.13 on page 72) and compare those results to the multispin system of a protein.

Figure 2.13 on page 72 showed that the cross-relaxation rate between two spins, σ, is positive for very small values of τ_c and has a zero-crossing at $\omega\tau_c \approx 1.12$ (corre-

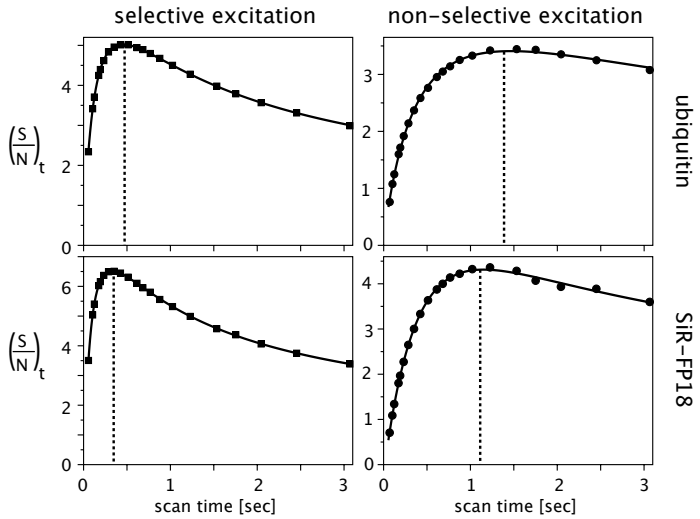


Figure 2.17 Experimental evidence for the τ_c dependence of the amide T_1 relaxation time. The curves show the sensitivity as a function of the scan time T_{scan} in an experiment using (left) selective 90° excitation of amide protons only (SOFAST-HMQC, see section 4) and (right) hard-excitation with water flip-back (se-wfb HSQC). Shown are data recorded at 25°C and 600 MHz ^{1}H frequency for ubiquitin (8.6 kDa, $\tau_c = 4$ ns, upper panels) and SiR-FP18 (E. coli sulfite reductase flavodoxine-like domain, PDB entry 1YKG, 18 kDa, $\tau_c \approx 8$ ns). These curves are adapted from figure 4.2 on page 107 and details can be found there.

sponding to $\approx$ 0.29 ns at 600 MHz ^{1}H frequency). For larger τ_c, the cross-relaxation rate is negative and its absolute value constantly increases. Because of this change in sign of σ, the selective relaxation of one spin is found to be *slowed down* for $\omega\tau_c <$ 1.12 relative to the non-selective case, whereas for $\omega\tau_c > 1.12$ an excited spin relaxes *faster* when its coupling partner is in equilibrium rather than excited (2.13(d, f)).

A similar behavior is found for selecive amide ^{1}H excitation in the multispin-system, shown in figure 2.16: As τ_c increases, the effective relaxation time found for the amide ^{1}H spins decreases. This can be attributed to the higher efficiency of spin-diffusion that "transports" the excitation energy to mobile sites where it gets efficiently evacuated (especially methyl groups). It is also interesting to note that the curves representing selective and non-selective excitation have a crossing point at low τ_c (central panel). For tumbling correlation times below this point the non-selective excitation yields *faster* recovery of the amide spins. This point is found at 0.29 ns ($\omega\tau_c$ = 1.12), as expected from the σ dependence.

The behavior of the relaxation times upon non-selective excitation is less intu-

itive. The effective average T_1 value has a minimum for very fast tumbling (at $\tau_c \approx$ 0.17 ns) and a maximum of T_1 is found for $\tau_c \approx 3$ ns. How can we interpret these results based on the observations made for the two-spin system ?

For the two-spin system a minimum of $T_1 = 1/(\sigma + \rho)$ upon non-selective excitation was found for $\tau_c \approx 0.17$ ns (at 600 MHz ^{1}H frequency, figure 2.13(b) on page 72). This corresponds exactly to the minimum found for the average amide ^{1}H T_1, indicating that at very small τ_c, where the cross-relaxation rate is very small the multi-spin system basically behaves like a collection of two-spin systems.

For higher τ_c values, T_1 in the two-spin system continuously increases with increasing overall tumbling correlation time, leaving open the question why in the multi-spin system a maximum and a subsequent decrease in T_1 is found. This effect can again be explained by cross relaxation as follows. Spin diffusion transports the energy efficiently to the sites where it gets evacuated, namely the mobile groups having high auto-relaxation rate constants. Note that the fast dynamics at these sites is retained irrespective of the overall tumbling correlation time. The higher τ_c the more efficient is this process.

There are therefore two counteracting processes: the increasing T_1 based on increasing $1/(\sigma + \rho)$ as predicted by the two-spin system on the one hand and the higher efficiency of spin diffusion together with local relaxation sinks on the other hand. As τ_c increases the latter effect outweighs the former and the balance of these processes therefore results in a maximum, which in principle can be expected to depend on the number and position of the mobile groups. Here a maximum is found in the range of a few nanoseconds. Many routine protein NMR studies happen to fall close to this range, which is from a proton longitudinal relaxation point of view the worst case in terms of sensitivity.

Finally, another factor can be studied, namely the effect of water exchange on the T_1 value in non-selective experiments. As discussed before, when water is in an excited state, exchange will result in slower magnetization recovery of protein-bound protons, which can be seen by comparing the dashed (water flip-back) and the dot-dashed (water-excite) curves in figure 2.16. The general behavior of T_1 as a function of τ_c is similar in the low-τ_c range, except for the absolute values of T_1, which is lower when water is in thermal equilibrium. However, for larger τ_c values an interesting feature is observed in the simulations: T_1 values *increase* again when τ_c exceeds some ≈ 50 ns. How can this be explained ?

In these simulations exchangeable sidechain protons were assumed to be in very fast exchange with water, and protons at these sites therefore relax with the relaxation rate of bulk water (3 seconds). As spin diffusion gets more and more efficient, the magnetization at the water-exchangeable sites is rapidly spread over the whole protein and in the extreme case it completely determines the relaxation

behavior of all protein-bound proton spins. The average effective amide ^{1}H T_1 is thus approaching the T_1 value of bulk water (here: 3 seconds). For very large proteins, a good water flip-back performance is therefore even more important than for smaller proteins, although for small proteins the losses are also substantial. Decreasing the bulk-water T_1 is another strategy that can be expected to be valuable, especially for large molecules (Hiller et al. (2005b)).

In summary, the τ_c dependence of T_1 yields insight into the mechanisms of proton relaxation in a multispin system. The relaxation behavior can be deduced from the one of the two-spin system taking into account spin diffusion and the differential auto-relaxation rates for different sites in the molecule. As spin diffusion gets very efficient, magnetization is efficiently "transported" to the sites where auto-relaxation is fast and where the energy of the spin system is evacuated efficiently.

Dependence on the magnetic field strength

The advent of higher magnetic field strengths is accompanied by several advantages in terms of resolution and sensitivity due to the higher Boltzmann populations and the larger frequency spread. What is generally less considered is the effect of higher field strength on the relaxation rates which in turn affects the sensitivity. Figure 2.18 shows the dependence of the average effective amide proton relaxation time constant in ubiquitin on the proton Larmor frequency in the case of selective amide excitation and non-selective excitation of protein protons while keeping water proton magnetization in equilibrium (water flip-back). Interestingly, in the case of non-selective excitation of protein proton spins, the relaxation rate continuously increases with increasing magnetic field strength. This finding highlights a feature of high magnetic fields that is rarely considered: although the equilibrium one-scan sensitivity increases due to the higher Boltzmann populations, longitudinal relaxation is slower and thus in multiscan experiments the sensitivity does not increase as much as one expects from the Boltzmann populations.

In the case of selective excitation the effective longitudinal relaxation time constant is almost independent of the field strength above ca. 200 MHz. The gain that selective excitation schemes offer thus increases with the magnetic field strength - making these schemes therefore more and more attractive. Note again a crossing point at a proton Larmor frequency of $\approx$ 45 MHz (which corresponds to $\omega\tau_c \approx 1.12$).

These simulation data on ubiquitin are corroborated by experimental data (figure 2.19). Whereas the optimal recovery delay (and thus the effective relaxation

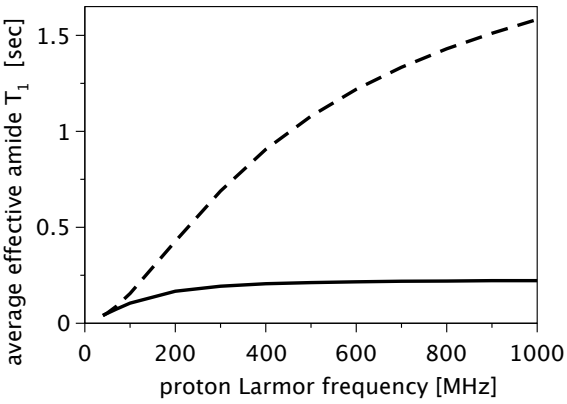


Figure 2.18 Simulated dependence of the effective relaxation time of the amide protons on the magnetic field strength. Shown are data for ubiquitin, the simulations were performed as before (τ_c=4ns, methyl rotation with τ_i=50 ps). The effective amide longitudinal relaxation rates in selective and non-selective experiments are shown by solid and dashed lines, respectively.

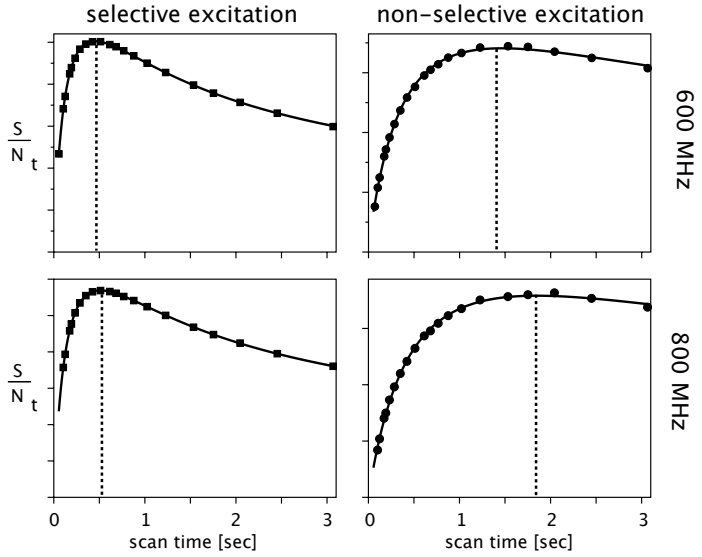


Figure 2.19 Experimental evidence for the B_0 dependence of amide relaxation rates. The curves show the sensitivity as a function of the scan time for ubiquitin at 600 and 800 MHz ^{1}H frequency. The maxima of the curves are related to the effective T_1 relaxation times. The data were acquired as in figure 2.17 and more details can be found there.

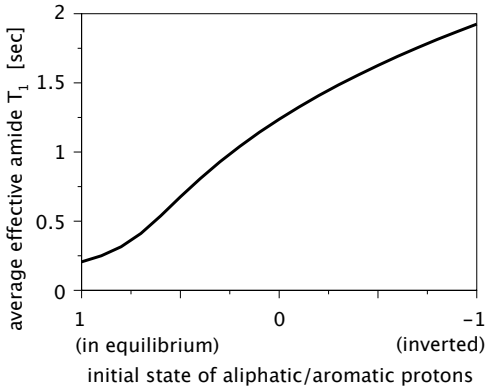


Figure 2.20 Simulated average proton longitudinal relaxation rate in ubiquitin as a function of the aliphatic/aromatic proton spin magnetization. 1 and -1 correspond to the aliphatic/aromatic magnetization vector aligned along +z and -z, respectively.

time) is almost independent of the magnetic field strength in the case of selective excitation, a clear shift to longer times when going to higher field is observed in the non-selective (water flip-back) experiment.

Amide relaxation and the initial magnetization of non-amide protons

So far only two extreme cases have been considered: either aliphatic and aromatic proton spins were fully aligned along +z (selective), or they were excited (i.e. in the transverse plane, non-selective). In practice, however, it might be difficult to keep the aliphatic/aromatic magnetization perfectly unperturbed when amide spins are excited and intermediate situations will be encountered. To estimate how much the amide relaxation will be altered in these intermediate situations, simulations were performed where the amount of aliphatic and aromatic magnetization is varied between a fully unperturbed state and a fully inverted state.

Figure 2.20 shows the average effective longitudinal relaxation time of amide protons in ubiquitin as a function of the initial state of the non-amide proton spin magnetization. These simulations show that fast amide relaxation critically relies on keeping the aliphatic/aromatic magnetization as closely as possible in an unperturbed state. Reducing the amount of magnetization on C-bound protons to half the maximum value yields a more than two-fold increase in the effective longitudinal relaxation. The slowest amide relaxation occurs in the case where the aliphatic spin magnetization is completely inverted yielding a more than 7-fold increase in the relaxation time.

Amide proton relaxation and internal dynamics

In all the above simulations, the protein was regarded as a rigid structure and the only internal dynamics that was considered was fast methyl rotation. This is of course an oversimplification and protein dynamics are omnipresent and often crucial for stability and function. Significant backbone and sidechain fluctuations are found especially in loop regions and terminal sequences, whereas the tightly packed secondary structure elements are often rather rigid.

The presence of such dynamics and structural heterogeneity has a number of consequences for amide proton relaxation that have hitherto been ignored.

First, internal dynamics change the spectral density function and therefore the auto- and cross-relaxation rates that govern the relaxation behavior (compare the corresponding expressions for methyl groups, equations 10.2 on page 215).

Second, dynamic (loop) regions are often located close to the protein surface in solvent-exposed regions of the protein. These regions are characterized by a less tight packing of atoms and therefore a smaller density of coupled spins. Additionally, amide protons in surface-exposed regions can undergo chemical exchange with water molecules.

All these factors alter the magnetization recovery and their exact nature is difficult to quantify in a simulation. To assess the role of internal dynamics on proton relaxation, experimental inversion-recovery data were analysed for amide ^{1}H spins in residues located in different structural elements. The behavior following non-selective and selective inversion is shown for three representative amide sites located in a rigid secondary structure, in a connecting loop, and at the C-terminus of ubiquitin (figure 2.21).

The residue located in a secondary structure (figure 2.21, Residue 4) senses the presence of many neighboring spins and the interactions are modulated basically by the overall tumbling time. This translates to a fast, bi-exponential recovery upon amide-selective inversion and a relatively slow recovery after non-selective inversion.

On the opposite extreme, the C-terminal amide proton (Residue 76) recovers its magnetization faster after non-selective inversion compared to amide H^N 4, but the recovery after non-selective excitation is slower, and does not show clear evidence for a bi-exponential behavior. Less effective spin diffusion and fewer coupling partners seem to be at the origin of this behavior.

Note that some structural and dynamic information is contained in the relaxation properties: large differences between the magnetization recovery following non-selective and selective inversion/excitation are found for amide sites located

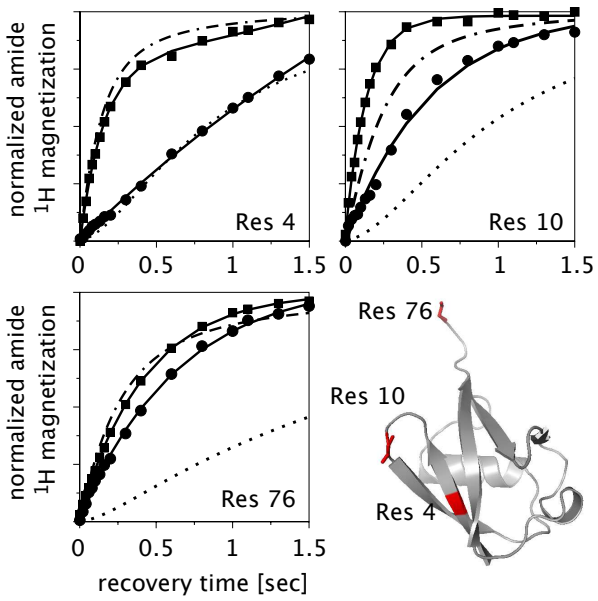


Figure 2.21 Experimental recovery curves of amide protons of Phe 4, Gly 10 and Gly 76 in ubiquitin following selective H^N inversion (filled squares) and non-selective inversion (filled circles). The solid lines are biexponential (monoexponential) fits to the experimental data for selective (non-selective) inversion. The dot-dashed and dotted lines represent simulations, with simulation parameters as used before. The locations of the three amide sites are indicated on the three-dimensional structure. The experimental conditions were as described in the caption of figure 2.14.

in rigid structures, whereas the differences are much smaller for exposed, dynamic amide sites. This property can actually be exploited to obtain a measure of structural compactness, as shown in chapter 8 (page 179).

The effect of deuteration on amide relaxation rates

The high density of proton spins in proteins does not only have consequences for the longitudinal relaxation but also for the transverse relaxation. In large protonated proteins, transverse relaxation rate constants are high, leading to severe line broadening and signal losses. This realization has led to the introduction of deuteration schemes that replace non-exchangeable protons by deuterons (Browne et al. (1973); LeMaster (1989, 1990a,b)).

Although deuteration proves a good method to decrease transverse relaxation rates (and therefore reduce signal losses), longitudinal relaxation is also slowed down.

Longitudinal relaxation-enhanced schemes can therefore be of special interest

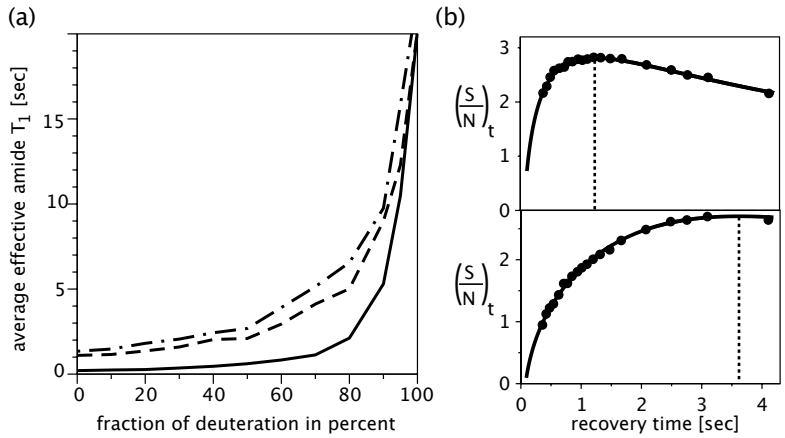


Figure 2.22 Effect of deuteration on the average effective T_1 values of amide protons in villin-HP36. (a) Simulation. Non-labile hydrogen sites were randomly protonated or deuterated corresponding to the displayed fraction of deuteration. Solid, dashed and dot-dashed lines correspond to selective, non-selective wfb, and non-selective non-wfb scenarios, respectively. For the latter case, water exchange and a bulk water T_1 value of 3 seconds was assumed, as before. The overall tumbling correlation time was set to 4 ns. (b) Experimental data of S/N_t on a sample of 75% deuterated SiR-FP18 for a selective experiment (upper panel) and an experiment that excites all protons (including water ^{1}H). As before, the maximum of these curves, indicated by dotted lines, is found at 1.25 T_1 (assuming monoexponential recovery). The data were measured with a BEST-Jcomp-HNHN-RDC experiment presented in chapter 6) and a non-selective version (Wu and Bax (2002)), also described there. These data are taken from figure 6.2 on page 146 and details can be found there. Note that the simulations were not performed with the same molecule (and τ_c) as the experiment, and a direct comparison is therefore not possible. However, the trend seen in the simulation is confirmed by experiment.

in such dilute proton spin systems. However, the selective excitation schemes illustrated above rely on the presence of a dipolar coupled network of spins, which raises the question whether they still are of benefit when applied to (partially) deuterated molecules.

Figure 2.22 (a) shows the average effective T_1 value for amide protons in villin-HP36 (in H_2O) as a function of the fraction of deuteration at non-exchangeable hydrogen sites. Here the percentage refers to the fraction of *non-labile* (generally carbon-bound) hydrogen sites that are occupied by ^{2}H. Exchangeable hydrogen sites were protonated. These curves predict that even for highly deuterated molecules substantially decreased amide ^{1}H relaxation times can be obtained upon selective excitation.

Spin diffusion in the network of remaining protons (all labile and part of the non-labile ^{1}H) still afford an efficient mechanism for cross-relaxation. As the fraction of deuteration approaches 100%, the recovery rate in selective and non-selective

experiments are approaching each other.

Figure 2.22 (b) displays experimental data found for ≈ 75% deuterated SiR-FP18 in H_2O for a selective experiment and a non-selective water-excite experiment. As above, the maxima of these curves are found at ≈ 1.25 times T_1. The experimentally found effective T_1, ≈ 1 second and 3 seconds, are in qualitative agreement with the simulations for SiR-FP18 at 75% deuteration (simulated data of SiR-FP18 not shown), although about 15-25% too high. The observed overestimation of T_1 may be due to the fact that other relaxation mechanisms (e.g. deuterium relaxation) as well as direct amide-water exchange was completely neglected in the simulation. Especially for very high degrees of deuteration, neglecting these effects should lead to an overestimation of the T_1 values.

In summary, even in highly deuterated molecules, where non-selective T_1 values can be on the order of several seconds, considerable acceleration of the amide relaxation rate is found by simulation and experiment for amide-selective experiments as compared to standard hard-excitation experiments.

Extension to a different group of protons: Longitudinal-relaxation optimized spectroscopy of aromatic rings

The concepts for enhanced longitudinal relaxation of amide proton spins introduced above can be extended to other spins. From the discussion about H^N spins, two requirements have to be met to obtain a significant effect: The number of protons that are excited in a selective-excitation experiment should be small compared to the spins left in equilibrium, and it must be experimentally feasible to solely excite the protons of interest with minimal perturbation of surrounding protons. In proteins two groups of protons, besides amide ^1H, appear to fulfill these criteria: aromatic protons and methyl protons.

Selective methyl experiments are shown in this thesis (chapter 4). Although some increase in their ^1H longitudinal relaxation is observed, the effects are relatively small for small to medium sized proteins, which can be assigned to the fact that their high mobility makes longitudinal relaxation fast already in the non-selective case. However, simulations show that for very large proteins the effect becomes more substantial (data not shown).

Aromatic protons are very promising candidates: They are relatively rare in most proteins and their resonance frequency is well separated from other proton spins, making it possible to manipulate them by selective pulses. Located in the hydrophobic core of proteins, they have a high proton-density in their environ-

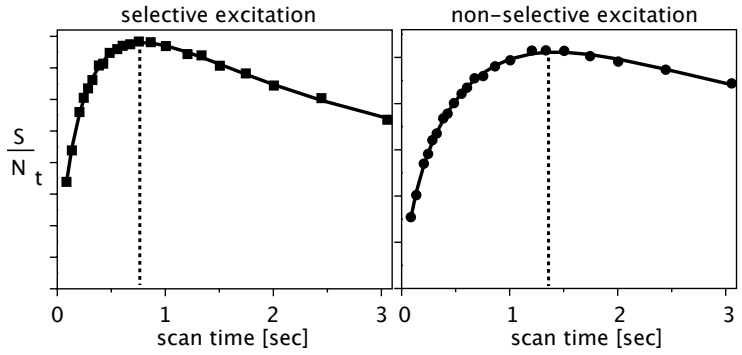


Figure 2.23 Longitudinal relaxation enhancement for aromatic protons. Sensitivity as a function of the scan time was measured for aromatic protons in ubiquitin (2 mM, 25°, 600 MHz ^{1}H frequency) using a standard se-wfb-HSQC sequence (right) and a SOFAST-HMQC experiment (left, see chapter 4). The displayed values are integrations of 1D spectra from 7 to 8 ppm. The se-wfb-HSQC experiment was used as implemented in the Varian BioPack, the SOFAST-HMQC experiment uses selective PC9 (Kupce and Freeman (1994)) and REBURP (Geen and Freeman (1991) pulses centered at 7.8 ppm with a band width of 3.5 ppm. The shift in the maximum of the curves indicates a $\sim$ two-fold reduction in the effective T_1.

ment and their internal dynamics on fast time scales is generally rather restricted; their location also makes them a valuable source of structural information.

Figure 2.23 shows the dependence of sensitivity as a function of the scan time in a selective experiment (left) and a non-selective experiment, demonstrating that the effective longitudinal relaxation time can be almost halved when using selective excitation.

Recently an experiment for resonance assignment of aromatic protons with flip-back of aliphatic proton spins has been proposed (Eletsky et al. (2005)).

Proton longitudinal relaxation in proteins: Summary and guidelines

The preceding pages have introduced the main features of proton longitudinal relaxation and it is worth summarizing the main conclusions that will also serve as guidelines for the design of relaxation-optimized fast-pulsing methods.

- The evolution of proton magnetization in proteins is governed by two main mechanisms: dipole-dipole interaction between the protons and exchange of labile hydrogen sites with proton spins from the solvent water.

- In compact, folded proteins, a simple model that ignores internal dynamics except fast methyl rotation reproduces experimental recovery data very well.

This points to the methyl groups as the main relaxation sink, while the slow overall tumbling allows for efficient spin-diffusion (see figure 2.14).

- Solvent exchange of labile protons introduces protons that are in the magnetization state of water protons into the protein. Therefore when water ^{1}H magnetization is in equilibrium, this process constitutes a pathway that evacuates excitation energy from the protein; in contrast, slowly-relaxing water ^{1}H spins that get incorporated to the protein slow down protein ^{1}H relaxation. Leaving water ^{1}H spins in an unperturbed state is therefore important for fast longitudinal relaxation of protein protons. The difference between water flip-back and water-excited experiments gets more important with increasing molecular weight due to the more efficient spin diffusion (see figure 2.16).

- Higher magnetic field strengths increase the effective proton relaxation times in non-selective experiments, whereas T_1 is almost independent of magnetic field strength, which makes these longitudinal relaxation optimized experiments particularly attractive at high fields.

- For rapid relaxation of amide proton spins, the unperturbed spins should be as close as possible to their thermal equilibrium magnetization. A deviation from the thermal equilibrium magnetization by only 50% already leads to a more than two-fold increase in the amide relaxation time (see figure 2.20).

- The proton relaxation upon selective and non-selective excitation depends on the structural context. In flexible regions of a protein the gain achieved with selective-excitation techniques is smaller (figure 2.21). The relaxation behavior of an amide proton after selective and non-selective excitation contains information about its structural environment (see also chapter 8).

- Even in highly (randomly) deuterated molecules amide proton relaxation can be enhanced using selective-excitation schemes.

The practical implementation of longitudinal relaxation enhanced experiments, guided by these conclusions, is shown in chapter 3

2.4 Summary

This chapter has introduced Fourier transform NMR, with a focus on its application to proteins. In order to resolve the signals of individual atoms, multidimensional techniques were found indispensable.

The canonical scheme for multidimensional NMR is intrinsically time consuming. In cases where the intrinsic sensitivity is low, this is not problematic, because in such cases multiple repetitions of the experiment have to be performed for reasons of signal accumulation. However, technical developments have greatly enhanced the sensitivity of many NMR experiments and often the duration of nD NMR experiments is no longer dictated by sensitivity considerations but by the need to sample the multidimensional data space; this situation is called *sampling limited regime*. Long experimental times are especially confining in situations where the sample stability is limited, in high-throughput NMR studies or when studying fast kinetic events.

This situation has triggered the development of different strategies to reduce the experimental time required to record multidimensional NMR spectra. The toolbox of methods is very rich today, and the approaches can be classified roughly in two groups: one aims at reducing the number of repetitions that are necessary to retrieve the desired spectral information, wheres the other tries to reduce the duration of one repetition.

Interestingly, these tools are often compatible with each other, and can be combined to yield optimized results in terms of sensitivity, speed and resolution for a given application. For example, the speed advantages of longitudinal relaxation optimized fast-pulsing approaches can be combined with reduced dimensionality, Hadamard, or spatial frequency encoding (Atreya and Szyperski (2004); Brutscher (2004); Gal et al. (2007); Schanda and Brutscher (2006); Schanda et al. (2006), see also chapter 7).

This chapter has introduced the principles underlying these approaches, and chapter 3 will focus on the practical implementation of longitudinal relaxation optimized NMR techniques for fast protein NMR.

3 Methods for longitudinal relaxation optimized protein NMR

Two promising spectroscopic methods for sensitive fast-pulsing NMR have been revealed in the preceding chapter: optimization of the effective flip angle (which results in a partial excitation) and selective excitation of a subgroup of protons (while leaving the other, unused protons in thermal equilibrium). Here, possibilities for the implementation of these two features are discussed, with a special focus on heteronuclear correlation experiments (^{1}H-^{15}N out-and-back techniques and related experiments). Some of these ideas were developed by other groups before or during my PhD thesis.

3.1 Review of experimental approaches

3.1.1 Implementation of Ernst angle excitation in nD experiments

Optimized flip angle excitation was initially proposed for the case of simple 1D 'pulse-acquire' experiments using a hard excitation pulse. The implementation of non-90° excitation in this case is achieved by simple reduction of the pulse length.

The situation becomes less straightforward for more complex pulse sequences used for multidimensional NMR. For example, a simple reduction of the pulse flip angle α of the initial pulse in an INEPT transfer block will *not* leave any proton magnetization aligned along +z: the concluding 90°(y) pulse of the INEPT block will convert the component $\cos\alpha H_z$ into $\cos\alpha H_x$, thus depleting the proton magnetization. Implementation of partial excitation in multidimensional NMR requires therefore a more careful design of pulse sequences.

In one case, however, Ernst angle excitation is still straightforward: In HMQC-

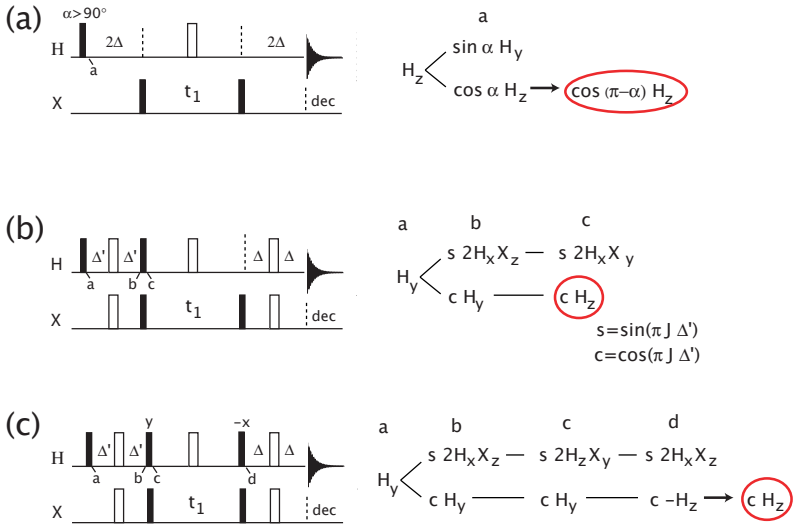


Figure 3.1 Implementation of Ernst-angle excitation in heteronculear H-X correlation experiments. (a) A simple adjustment of the initial excitation pulse angle α can be done in HMQC type experiments. This results in a longitudinal magnetization of $\cos\alpha$, and after the 180° refocusion pulse the final longitudinal magnetization is given by $\cos(180° - \alpha)$. (b) Partial excitation in a HMQC experiment using a J coupling delay $< 1/(2J)$, resulting in a remaining H_z term at point c, which is retained by the two following π rotations. (c) HSQC implementation of the same idea as in (b). For simplicity, the gradient pulses and phase settings for quadrature detection are omitted in the pulse schemes.

type experiments applying a pulse with flip angle $\alpha > 90°$ results in a partial restoration of proton magnetization after the subsequent π pulse. Figure 3.1(a) shows a Ernst-angle HMQC sequence, called Fast-HMQC, proposed by Ross et al. (1997). The authors proposed the experiment for fast ligand screening and demonstrated gains of $\sim$ 25 % when using this optimized flip-angle Fast-HMQC pulse sequence relative to a comparable HMQC sequence employing 90° excitation. [1]

A different way of implementing a partial excitation is to use a 90° excitation pulse, and then separate the coherence in two components, and flip back one of the components, thereby achieving a net partial excitation.

For example, a H_y coherence of a proton coupled to a heteronucleus will evolve during a delay Δ' as $\sin(\pi J\Delta')\ 2H_xN_z + \cos(\pi J\Delta')\ H_y$. Setting the delay Δ' to a

[1] Note, however, that a comparison to the standard pulse sequence used for the same purpose, sensitivity-enhanced HSQC (Kay et al. (1992)), reveals that it performs actually less well than se-HSQC, except for short recovery delays $T_{rec} \leq 100$ ms, where it becomes comparable to its HSQC counterpart (Schanda and Brutscher (2005)). This can be attributed to the difference of the quadrature detection schemes in the two experiments.

value shorter than $1/(2J)$ leaves a non-zero H_y component, which can be manipulated independently from the orthogonal $2H_xN_z$.

Kupce and Freeman (2007) have shown this concept for a ^{1}H-^{13}C HMQC-type experiment (figure 3.1b). The component H_y that results from incomplete J coupling evoution can be converted to H_z by a 90° ^{1}H pulse, thereby resulting in the desired partial excitation. The $2H_xC_z$ coherence is not affected by this pulse, and ultimately gives rise to the detected signal.

This experiment has an additional nice feature: not only is a part of the magnetization of the proton spins of interest (coupled to X) restored, but the magnetization of all the non-X-coupled spins is also retained in equilibrium, and this magnetization can then be used to accelerate the recovery of the magnetization of the used ^{1}H spins. The experiment was proposed for natural abundance samples (for ^{1}H-^{13}C correlation experiments), where all the magnetization of ^{12}C bound proton spins is effectively restored before acquisition.

In principle, a similar idea can be used also for HSQC experiments, although this has not yet been demonstrated in the literature. Figure 3.1(c) shows how such a Ernst-angle HSQC sequence may be designed. As in the preceding HMQC example, incomplete J-coupling evolution is used to create $2H_xN_z$ and H_y coherences. In this case, however, the component that did not evolve under the coupling is spin-locked in the transverse plane by the second 90° pulse. After t_1 evolution (and maybe additional transfer steps), the concluding reINEPT block transforms the H_y component back into H_z, resulting thus in a partial flip-back.

The advantage of such a sequence is that it creates single-quantum ^{15}N coherence, which is favorable in terms of transverse relaxation as compared to multiple-quantum evolution (see e.g. Cavanagh et al. (1995), page 429). This method could in principle be extended to triple-resonance experiments. The drawback of this technique comes from the fact that the magnetization, which is ultimately flipped back, is stored as H_y during the sequence. Transverse relaxation losses during the t_1 evolution can result in decreased flip-back performance, especially for large molecules. Keeping in mind that the magnetization that is flipped back was initially separated from the one that actually gives rise to the detected signal, this partial excitation approach may not yield higher sensitivity than a 90° excitation with properly set delays $\Delta' = \Delta = 1/(2J)$.

To conclude, partial excitation in multidimensional NMR experiments is not straightforward, but may be implemented in some selected cases. Especially, a simple adjustment of the excitation angle can be implemented in HMQC-type experiments. As this approach does not rely on the intermediate creation of proton coherence that is then converted to H_z, it seems to be the most promising approach

for large molecules, such as proteins. We therefore focus on this strategy (see below).

3.1.2 Selective excitation of a subgroup of protein proton spins

Separating the observed spins, that follow the desired coherence pathway and ultimately give rise to the signal, from the unused spins, that are kept close to their equilibrium state, is at the heart of longitudinal relaxation optimized (L-opt) protein NMR. L-opt experiments have to meet two criteria to yield high sensitivity: First, as in any experiment, the main coherence transfer pathway, which determines the type of information that is obtained, should be as efficient as possible. Relaxation losses and pulse imperfections should be as small as possible. Second, the flip-back of the unused proton spins should be as "clean" as possible to obtain an optimal longitudinal relaxation enhancement on the observed proton spins. Only when the unused spins are close to thermal equilibrium maximal longitudinal relaxation enhancement of the observed spins is achieved (see figure 2.20 on page 87). In the following we consider the amide ^{1}H spins as the actually observed spins.

Selective manipulation of the different subgroups can be based on different physical principles: either the different J coupling topology e.g. of ^{15}N and ^{13}C bound protons is exploited, or the frequency separation between used and unused spins is exploited by application of selective pulses.

3.1.2.1 Separation by scalar couplings

Those methods that exploit J couplings to separate used from unused spins, can again be divided in two groups:

The first group achieves a separation by the coupling that is exploited in the main coherence transfer pathway (i.e. the $^{1}J_{NH}$ coupling in ^{1}H-^{15}N experiments). This results in a uniform flip-back of all not-^{15}N bound proton spins and is referred to as *passive separation*. An example of such a technique was recently proposed by Diercks and co-workers (Diercks et al. (2005)) for HNCX (CX=CO,CA) experiments. In principle, this approach should give perfect flipback of all unused spins. In practice, however, this is not observed: only 30 to 50% of the aliphatic magnetization could be restored back by this scheme (Diercks et al. (2005)), which yields only minor amide relaxation enhancements (see figure 2.20 on page 87). Transverse relaxation losses of the non-^{15}N bound spins, as well as pulse imperfections

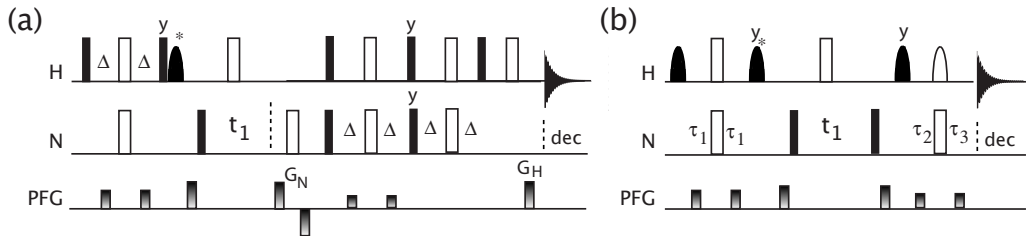


Figure 3.2 Selective HSQC experiments proposed previously. (a) Simply replacing the water flip-back pulse by a selective pulse covering aliphatic and water protons is the simplest method to achieve (partial) flip-back of these protons. Filled and open symbols denote 90° and 180° pulses, and the shaped pulse is a EBURP* flip-back pulse (time reversed EBURP) covering the range -1 to 5 ppm. Coherence selection by gradients is performed as described elsewhere (Kay et al. (1992)). Delays Δ are set to $1/(4J_{NH})$. This is the aliphatic flip-back pulse sequence used for collecting the data shown in figure 4.3(c) on page 109. (b) Longitudinal relaxation optimized (L-opt) HSQC sequence proposed by Pervushin et al. (2002). Filled and open selective pulses are applied with a EBURP, and REBURP shape, respectively, covering a bandwidth of 4 ppm; the pulse indicated with an asterisk is a EBURP* flip-back pulse. The delay τ_1 is set to 2.4 ms, τ_2, τ_3 are set to a shorter value (ca. 1 ms at 600 MHz) to account for coupling evolution during the REBURP pulse. This sequence was used to record the data shown in figure 4.2 and 4.3(b).

were identified as the main sources for aliphatic magnetization loss.

The second approach uses actively the scalar coupling of the unused carbon-bound ^{1}H spins to create $2H_zC_z$ spin order, while the ^{15}N bound proton spins evolve to $2H_zN_x$. The (slowly relaxing) two-spin order $2H_zC_z$ is then converted to H_z by the following INEPT blocks. Examples were proposed for TROSY-type H-N correlation experiments (Pervushin et al. (2002)) and HNCO-type experiments (Diercks et al. (2005)). The advantage of this scheme is that the main coherence pathway is not affected: in the simplest case, only ^{13}C pulses have to be inserted in the ^{1}H-^{15}N INEPT blocks.

However, transverse relaxation losses, the non-uniformity of ^{1}H-^{13}C coupling constants and pulse imperfections lead to significant reduction of the aliphatic magnetization that gets restored at the end of the pulse sequences, and only negligibly small sensitivity advantages are reported by Diercks et al. (2005).

3.1.2.2 Separtion by selective pulses

Based on their good spectral separation (see figure 2.4), amide or aromatic protons can be manipulated independently from the other proton spins. Here, one can discriminate between experiments that use selective pulses either for the unused or

for the observed protons. Figure 3.2(a) shows the simplest example of a ^{1}H-^{15}N experiment with aliphatic flip-back, where a selective pulse is applied to the aliphatic and H_2O signals after the initial INEPT transfer block (replacing the water flip-back pulse often applied at this place). The remaining part of the sequence then results in a net 720° rotation, restoring longitudinal magnetization ^{1}H spins resonating in the frequency range covered by the selective pulse. A related experiment for aromatic HCCH correlation spectra has been reported by Eletsky et al. (2005).

The drawback of these approaches is that the unused protons go through long coherence transfer pathways, spending thus much time in the transverse plane, which leads to less-than-optimal flip-back performance because of T_2 losses.

Diercks and co-workers have therefore proposed a HNCO/CA experiment, which applies with *each* hard proton 90° pulse a band-selective aliphatic flip-back pulse. The ratio of aliphatic flip-back that can be achieved with this scheme is about 50 to 60% in the slow pulsing limit (Diercks et al. (2005), see section 5.3). The problem of this approach is that the aliphatic proton spins are manipulated by a large number of selective and non-selective pulses, and pulse imperfections and B_1 inhomogeneity can rapidly deterior the performance of this scheme.

Selective manipulation of the used proton spins, rather than flip-back of the unused spins, is an alternative (Deschamps and Campbell (2006); Pervushin et al. (2002)). The pulse sequence in figure 3.2(b) is one way to implement an amide-proton selective HSQC. It does not rely on separation by scalar coupling evolution, and the aliphatic proton spins do not spend any time in the transverse plane (the flip-back performance is therefore not deteriorated by transverse relaxation losses).

The drawback of methods that rely on selective pulses is that signals outside the chosen pulse bandwidth, or close to the edge, may be absent in the spectrum, or may have reduced intensity. In addition, the efficiency of coherence transfer on the main coherence transfer pathway may be reduced when using selective pulses, as compared to the use of hard pulses. A careful choice of selective pulses that have a good performance within the excitation bandwidth and a clean off-resonance behavior is therefore crucial for these methods.

In terms of flip-back performance, we find that the selective manipulation of the used spins yields the best results (see section 5.3). We have therefore focussed on methods using selective amide ^{1}H pulses.

4 SOFAST-HMQC experiments for recording two dimensional heteronuclear correlation spectra of proteins within a few seconds

4.1 Introduction

^{1}H-^{15}N correlation spectra serve as a fingerprint of proteins by NMR spectroscopy: yielding one correlation peak per amino acid, they often suffice to resolve most of the signals corresponding to the individual residues along the polypeptide chain. Their amino-acid resolved information makes such spectra generally the first step in every protein study by NMR. In principle, ^{1}H-^{15}N correlation maps are very useful either to screen many different samples (or sample conditions) or to follow spectral changes, e.g. during a kinetic reaction. However, standard methods generally require experimental times on the order of minutes, which excludes many interesting phenomena in proteins occurring on a seconds time-scale from site-resolved studies.

In an attempt to speed up data acquisition for site-resolved NMR, we introduce here a new experiment for very fast data acquistion. It makes use of Ernst angle excitation combined with longitudinal relaxation optimization. We show that 2D correlation spectra of proteins at millimolar concentrations can be acquired in several seconds. The proposed experiment proves very useful not only for very rapid data acquisition, but we also show that it is the most sensitive technique compared to previously proposed experiments, such as se-wfb HSQC (Kay et al. (1992)).

4.2 SOFAST-HMQC experiments

4.2.1 General features of SOFAST-HMQC

The pulse sequences used to record band-Selective Optimized-Flip-Angle Short-Transient (SOFAST) HMQC spectra are shown in figure 4.1. These pulse sequences provide the required high sensitivity to perform fast heteronuclear H-X correlation experiments of macromolecules by using very short recycle delays (t_{rec}). The main features of SOFAST-HMQC are the following: (i) the HMQC-type H-X transfer steps require only few rf pulses which limits signal loss due to B_1-field inhomogeneities and pulse imperfections. A reduced number of rf pulses will be especially important if the experiment is performed on a cryogenic probe, where B_1-field inhomogeneities are more pronounced. (ii) The band-selective ^{1}H pulses reduce the effective spin-lattice relaxation times (T_1) of the observed ^{1}H spins. The presence of a large number of non-perturbed ^{1}H spins, interacting with the observed ^{1}H via dipolar interactions (NOE effect), significantly reduces longitudinal relaxation times whereby the equilibrium spin polarization is more quickly restored. As will be shown later on, the longitudinal relaxation optimization enhancement effect depends on the number and type of the applied ^{1}H pulses. The use of only 2 (band-selective) ^{1}H pulses in SOFAST-HMQC ensures minimal perturbation of the undetected proton spins, and provides higher enhancement factors than observed with other longitudinal relaxation optimized pulse schemes (Pervushin et al. (2002)). As long as the water resonance is outside the selected ^{1}H pulse bandwidth, the WATERGATE-type (Piotto et al. (1992)) pulse sequence element G_1-180° (^{1}H)-G_1 yields efficient water suppression within a single scan. The selective ^{1}H manipulation also removes coupling evolution between excited ^{1}H spins and passive ^{1}H spins from frequency bands that are not perturbed by the selective pulses. (iii) The adjustable flip angle (Ernst angle) of the ^{1}H excitation pulse allows further enhancement of the available steady-state magnetization for a given recycle delay. Ernst-angle excitation in HMQC-type correlation experiments has been proposed previously (Ross et al. (1997), figure 3.1(a)). With respect to this Fast-HMQC, our new SOFAST-HMQC experiment has several advantages: Because of the selective ^{1}H manipulations, ^{1}H spin-lattice relaxation times are significantly reduced, and ^{1}H-^{1}H coupling evolution during t_1 and the transfer delays Δ is refocused. An additional $\sqrt{2}$ sensitivity gain is achieved because of the amplitude-modulated instead of phase-modulated quadrature detection scheme, required in the non-selective Fast-HMQC experiment of Ross et al. for efficient solvent suppression in aqueous protein samples. As only one of the 2 orthogonal components present after the ^{15}N frequency labeling (t_1) is transferred into observable ^{1}H coherence by the Fast-HMQC sequence, this phase-modulated quadrature

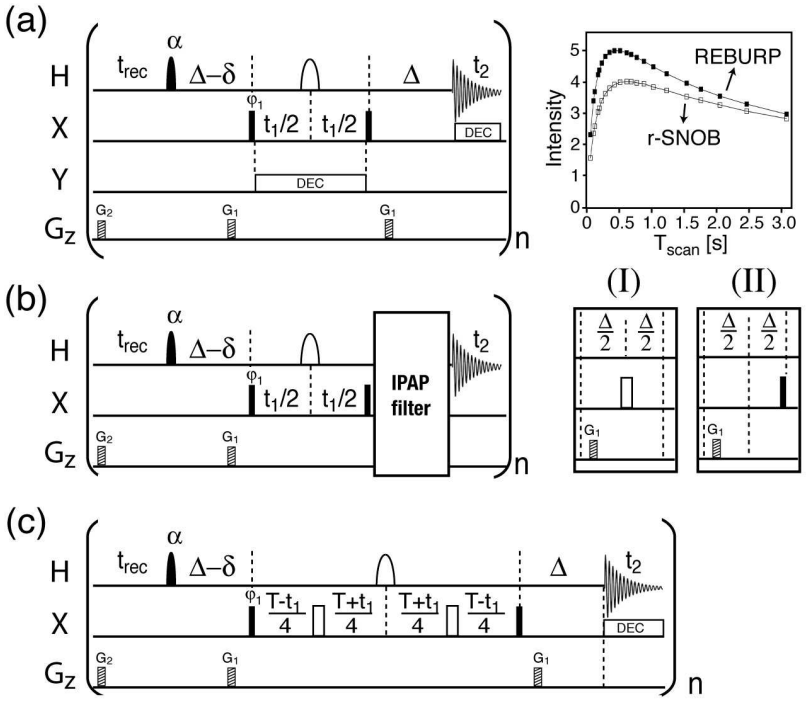


Figure 4.1 Pulse sequences used to record SOFAST-HMQC ^{1}H-X (X=^{13}C or ^{15}N) correlation spectra: (a) basic pulse scheme, (b) IPAP version, and (c) CT version. Filled and open pulse symbols indicate 90° and 180° rf pulses, except for the ^{1}H excitation pulse applied with flip angle α. The variable-flip-angle pulse has a polychromatic PC9 shape (Kupce and Freeman (1994)), and band-selective ^{1}H refocusing is realized using either REBURP (Geen and Freeman (1991)) or r-SNOB (Kupce et al. (1995)). The transfer delay Δ is set to $1/(2J_{HX})$, and t_{rec} is the recycle delay between scans. The delay δ accounts for spin evolution during the PC9 pulse, and has to be adjusted prior to data acquisition to yield pure-phase spectra in the ^{1}H dimension. The CT delay T in (c) is set to $T = n/J_{XX}$ with n an integer. Quadrature detection in t_1 is obtained by phase incrementation of φ_1 according to STATES or TPPI-STATES. In standard SOFAST-HMQC (a) adiabatic WURST-2 decoupling (Kupce and Wagner (1995)) is applied on X during detection. For IPAP-SOFAST-HMQC two experiments are recorded as indicated in the inserts of (b). In experiment (I), ^{1}H-X coupling evolution during Δ is refocused by application of a 180° X pulse. In experiment (II), ^{1}H-X coupling evolution is active during Δ. The 90° X pulse applied before final detection converts any remaining antiphase coherence of the type $2H_xX_z$ into undetectable multiple quantum coherence. This purge pulse is mainly useful for the measurement of H-X coupling constants of partially aligned protein samples. The two data sets (I) and (II) are then added or subtracted prior to Fourier transformation to yield the upfield or downfield components of the ^{1}H doublet line, respectively. Note that IPAP filtering can also be applied in the CT-SOFAST-HMQC sequence shown in (c) by simply omitting the X decoupling during detection, and replacing the back-transfer delay Δ by the two inserts shown in (b). The graph in the upper right corner shows a comparison of the performance of REBURP (upper curve) and r-SNOB (lower curve) when used in the SOFAST-HMQC sequence (a) to record ^{1}H-^{15}N correlation spectra. The data were recorded at 600 MHz ^{1}H frequency on a sample of ^{15}N-labeled ubiquitin using a flip angle α=90°.

detection scheme results in a $\sqrt{2}$-sensitivity loss relative to amplitude-modulated quadrature detection used in SOFAST-HMQC (Schanda and Brutscher (2005)).

4.2.2 Band-selective ^{1}H pulses in SOFAST-HMQC

The performance of SOFAST-HMQC critically depends on the choice of the pulse shapes for the band-selective excitation and refocusing pulses on the ^{1}H channel. We have tested several pulses to achieve excitation and refocusing with minimal perturbation of aliphatic proton spins and minimal B_1 inhomogeneity losses.

For refocusing we initially chose a r-SNOB profile (Kupce et al. (1995)), which presents the advantage of a short pulse length thus reducing signal loss due to transverse spin relaxation. However, we find that for the recording of ^{1}H-^{15}N correlation spectra, a REBURP profile, although of a three times longer duration, yields higher sensitivity. An experimental comparison of r-SNOB and REBURP performance in ^{1}H-^{15}N SOFAST-HMQC is shown in figure 4.1(a). A signal increase of up to 50% is observed when using REBURP instead of r-SNOB for short scan times. This surprising result can be explained by the better off-resonance performance of REBURP, resulting in less perturbation of the aliphatic ^{1}H spin polarization and, as a consequence, shorter longitudinal relaxation times of the amide proton spins. This is manifest by the maximum of the intensity curves, shifted towards shorter scan times for the experiments realized using REBURP. A similar behavior was also observed for other proteins varying in size and different magnetic field strengths (data not shown).

In order to implement the feature of optimized flip-angle band-selective excitation in the SOFAST-HMQC we have explored the literature for suitable pulse shapes. Most of the band-selective top-hat pulse shapes commonly used for NMR spectroscopy, e.g. BURP (Geen and Freeman (1991)), Gaussian pulse cascades (Emsley and Bodenhausen (1992)), or SNOB (Kupce et al. (1995)), have only been optimized for discrete flip angles of 90° or 180°, and generally are not useful for variable flip angle excitation purposes. In contrast, polychromatic (PC) selective pulses have been shown to perform well for a whole range of flip angles (Kupce and Freeman (1994)). These PC pulses are based on a series of simultaneously applied, frequency shifted basic pulse elements. For the present applications we have used the PC9 pulse shape, which has the desired "top-hat" excitation profile for flip angles $0° < \alpha <$120°. For the flip-angles of 120° to 150° used in this work, the excitation profile slightly deteriorates, but the PC9 pulse shape still yields good experimental results. For the 90° to 150° flip-angles used in this work the combination of PC9 and REBURP (or r-SNOB) pulses was found a robust choice. Unlike other band-selective excitation pulses that yield "pure-phase" transverse magnetization, the PC9 pulses produce phase that is a linear function of the frequency

offset. Schematically, one can thus replace a PC9 pulse by the combination of a pure-phase excitation pulse followed by a delay δ. The chemical shift and scalar J_{HX} coupling evolution occurring during this delay δ can be accounted for by adjusting the subsequent transfer delay of the HMQC sequence to $1/(2J_{HX}) - \delta$ (see Figure 4.1). If the delay Δ has been properly adjusted prior to data acquisition no first-order phase correction is required in the ^{1}H dimension. Otherwise, pure-phase spectra can still be obtained by applying a first order phase correction.

4.2.3 IPAP-SOFAST-HMQC and CT-SOFAST-HMQC

Alternatives to the basic SOFAST-HMQC pulse scheme are shown in Figs. 4.1(b) and c. In IPAP-SOFAST-HMQC (Figure 4.1b) the heteronuclear X decoupling during acquisition (t_2) is replaced by an IPAP filter (Andersson et al. (1998); Ottiger et al. (1998)). The IPAP filter is realized by the two sequence blocks shown in the inserts of figure 4.1b. Two spectra have to be recorded with and without refocusing of the ^{1}H-X coupling evolution during the back-transfer delay Δ, thus increasing the minimal experimental time by a factor of 2. Addition or subtraction of these spectra then separates the two ^{1}H doublet components in different sub-spectra. Each of the 2 sub-spectra has a 2-times lower S/N ratio than a standard SOFAST-HMQC spectrum recorded in the same experimental time (corresponding to the acquisition of both sub spectra). A single correlation spectrum is obtained by adding the 2 sub-spectra after appropriate shifting of the two spectra with respect to each other in the ^{1}H dimension by an amount of $\pm(J_{HX})/2$. This results in a $\sqrt{2}$ increase in S/N ratio, and consequently the sensitivity (S/N per unit experimental time) of IPAP-SOFAST-HMQC is decreased by a factor $\sqrt{2}$ with respect to the standard SOFAST-HMQC pulse scheme of Figure 4.1a. For larger molecules and high magnetic field strengths cross-correlated relaxation effects induce differential broadening of the ^{1}H doublet lines in the absence of ^{15}N decoupling. Adding the 2 lines, a broad and a narrow one, will result in complicated (non Lorentzian) line shapes in the ^{1}H dimension, but also increases the sensitivity of IPAP-SOFAST-HMQC (TROSY effect). The main interest of the IPAP-version of SOFAST-HMQC is that it does not require composite X decoupling during detection. Reducing rf power becomes an important issue when applying the SOFAST-HMQC experiment on NMR spectrometers equipped with a cryogenic probe. In this case the $\sqrt{2}$-sensitivity loss due to the IPAP filtering is largely compensated by the 2 to 4-times higher sensitivity provided by the probe. IPAP-SOFAST-HMQC also offers the possibility of fast measurement of heteronuclear H-X (scalar or residual dipolar) spin-spin coupling constants yielding useful additional probes of structural changes occurring during kinetic processes that can be monitored by real-time NMR. SOFAST-HMQC can also be combined with constant time (CT) frequency

editing in the indirect dimension (t_1). The pulse sequence for CT-SOFAST-HMQC is shown in Figure 4.1c. This CT version is mainly useful for removing line splittings due to ^{13}C-^{13}C couplings in ^{1}H-^{13}C correlation experiments, thus providing increased spectral resolution at the expense of sensitivity. The experiment can be either implemented with heteronuclear decoupling during detection or an IPAP filter for application on cryogenic probes, or measurement of ^{1}H-X spin coupling constants.

4.3 Application to proteins

4.3.1 ^{1}H-^{15}N SOFAST-HMQC

The SOFAST-HMQC pulse sequences of figure 4.1 have been designed to provide high sensitivity for fast repetition rates. To examine the performance of the SOFAST-HMQC experiment for the desired short interscan delays we have measured 1D spectra of ^{15}N-labeled ubiquitin (7 kDa, 2 mM, 25°C, pH 6.2) and of ^{13}C,^{15}N-labeled SiR-FP18 (18 kDa, 1.8 mM, 25°C, pH 7.0), the flavodoxin-like domain of the *E. coli* sulfate reductase, at magnetic field strengths corresponding to 600 and 800 MHz ^{1}H frequency using the pulse sequence of figure 4.1(a) without ^{15}N decoupling during detection, and setting t_1=0. The 800 MHz spectrometer was equipped with a cryogenic probe, whereas the experiments at 600 MHz ^{1}H frequency were performed on a standard (non cryogenic) probe. Figure 4.2 shows the measured S/N ratios for constant experimental time as a function of the duration of a single repetition of the experiment T_{scan} (taking into account the length of the pulse sequence, data acquisition time, and recycle delay) for ubiquitin (figs. 4.2a and b) and for SiR-FP18 (figs. 4.2c and d). Each intensity point was obtained by scaling all spectra to the same noise level according to the number of applied scans, and integrating the spectral intensity over the range 7.0 to 9.5ppm. The curves are therefore representative of the average behavior of the experiment for all amide sites in the protein. The SOFAST-HMQC data for three different flip angles (90°, 120°, and 150°) are compared to results from a sensitivity-enhanced (se) water-flipback (wfb) HSQC pulse sequence as implemented in the Varian Bio-pack, and from a longitudinal relaxation optimized HSQC (LHSQC) experiment (Pervushin et al. (2002)). The peak intensities in the se-wfb-HSQC spectra were up-scaled by a factor of $\sqrt{2}$ to account for the sensitivity-enhanced quadrature detection providing a $\sqrt{2}$-signal enhancement in the 2D version of the experiment with respect to standard quadrature detection as implemented in SOFAST-HMQC and LHSQC. Theoretically we expect that the different quadrature detection schemes translate into a sensitivity advantage of $\sqrt{2}$ for se-wfb-HSQC with

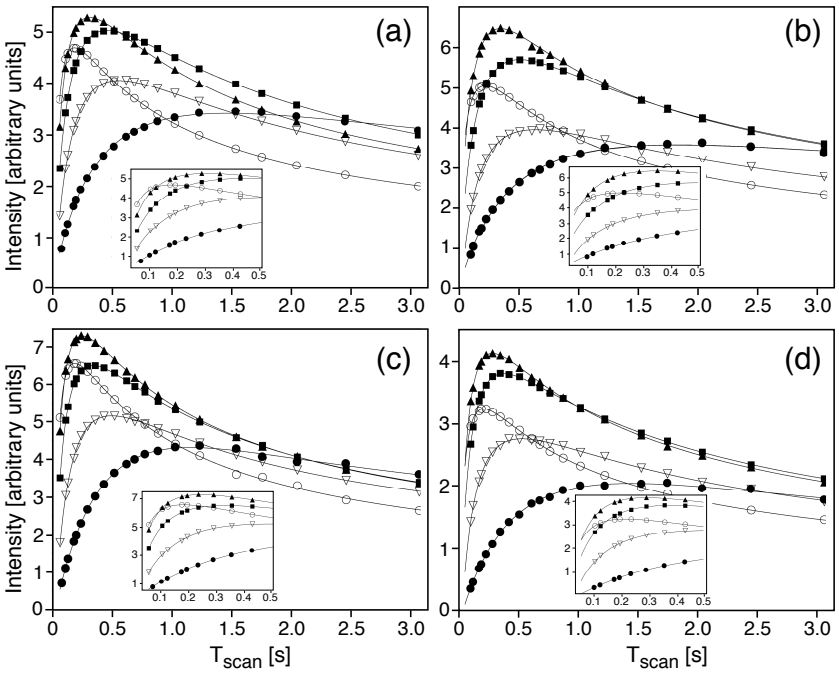


Figure 4.2 Sensitivity plotted as a function of the scan time (T_{scan}) obtained with different ^{1}H-^{15}N correlation experiments for (a) ubiquitin (6.8 kDa) at 600MHz, (b) ubiquitin at 800 MHz, (c) SiR-FP18 (18 kDa) at 600 MHz, and (d) SiR-FP18 at 800 MHz. The 800 MHz spectrometer was equipped with a cryogenic probe. The intensities were extracted from 1D spectra recorded using the SOFAST-HMQC sequence of fig. 4.1a (t_1=0) with flip angles of α = 90° (squares), 120° (triangles) and 150° (open circles), LHSQC (open triangles), and se-wfb HSQC (filled circles). Band-selective ^{1}H pulses in the SOFAST-HMQC and LHSQC experiments were centered at 8.0 ppm covering a bandwidth of 4.0 ppm. Variable flip angle excitation and refocusing in SOFAST-HMQC were realized using PC9 and REBURP pulse shapes, respectively. The se-wfb-HSQC experiments were recorded using the pulse sequence implemented in the Varian Bio-pack. The LHSQC sequence was set up as described by Pervushin et al. (2002), except for the final INEPT delays τ_2 and τ_3 that were reduced to 1.0 ms to account for spin evolution during the REBURP pulse. The inserts show an expansion of the data for short scan times. Each spectrum was acquired in the same experimental time, and the spectra were scaled to the same noise level. The se-wfb-HSQC spectra were scaled up by a factor $\sqrt{2}$ to account for the gain in S/N ratio obtained by the sensitivity-enhanced quadrature detection in the 2D version of the experiment. Intensities were obtained by integration of the 1D ^{1}H spectra over the spectral range 7.0 to 9.5 ppm. The solid lines are smoothened interpolations of the experimental data points. Fitting the data points to equation 2.27 shows systematic deviations, which is explained by the non-monoexponential behavior (see the builup curves in figure 2.14 on page 77).

respect to SOFAST-HMQC and LHSQC at long recycle delays. In practice, we find that this sensitivity difference is partly or completely compensated by the larger number of pulses applied in se-wfb-HSQC leading to increased signal losses due to pulse imperfections and B_1-field inhomogeneities. As expected, this effect is more pronounced for the data acquired using a cryogenic probe (Figs. 4.2(b) and (d)), characterized by larger B_1-field gradients across the detection volume. For shorter scan times the reduced spin-lattice ^{1}H relaxation times become important, resulting in a higher sensitivity for the longitudinal-relaxation optimized pulse sequences (SOFAST-HMQC and LHSQC). The more efficient spin-lattice relaxation is reflected in the shift of the maximum of the intensity curves (T^{opt}) towards shorter scan times. This shift is more pronounced for SOFAST-HMQC than for LHSQC. This observation is most likely explained by the larger number of ^{1}H pulses required for LHSQC, resulting in some perturbation of the aliphatic and water ^{1}H equilibrium spin polarization. Using the relation T^{opt} = 1.25·T_1 (see eq. 2.28 on page 64) one can estimate average amide ^{1}H spin-lattice relaxation times of $T_1 \approx$ 0.9-1.4 s for HSQC, of $T_1 \approx$ 0.4-0.5 s for LHSQC, and of $T_1 \approx$ 0.3-0.4 s for SOFAST-HMQC (α = 90°). For all experiments, we observe an increase in the average T_1 times with increasing B_0 field strength and with decreasing molecular weight (or tumbling correlation time) of the molecule (see simulations in section 2.3.2.2, pages 82 and 86). For very short scan times the sensitivity advantage of SOFAST-HMQC with respect to se-wfb HSQC and LHSQC becomes even more pronounced when using larger flip angles α >90°. As shown in the inserts of figs. 4.2a-d important signal enhancements are observed for $T_{scan} < 200$ ms when using SOFAST-HMQC instead of se-wfb HSQC or LHSQC. This sensitivity enhancement allows significantly reduced experimental times for a given S/N ratio.

The 1D spectra in Figure 4.2 provide only information on the average signal to noise ratio obtained by the different pulse sequences. In order to look in more detail at the spread of peak intensities among the amide protons in SOFAST-HMQC, we have recorded a 2D ^{1}H-^{15}N correlation map of ubiquitin at 600 MHz ^{1}H frequency using a short scan time of T_{scan}=165 ms and an optimized flip angle α=120° (see figure 4.2a). This spectrum, shown in Figure 4.3a can be compared to spectra recorded using se-wfb-HSQC (Fig. 4.3d), LHSQC (Fig. 4.3b), and sensitivity-enhanced aliphatic-flip-back HSQC (se-afb-HSQC) pulse sequences (Fig. 4.3c). For se-afb-HSQC the water flip-back pulse in the standard se-wfb-HSQC sequence has been replaced by a band-selective EBURP flip-back pulse covering the ^{1}H frequency range from -1 to 5 ppm. Although, for this scan time all longitudinal-relaxation optimized experiments provide a significant gain in signal to noise ratio with respect to the standard se-wfb-HSQC experiment, SOFAST-HMQC is by far the most sensitive. A residue-by-residue analysis of the peak

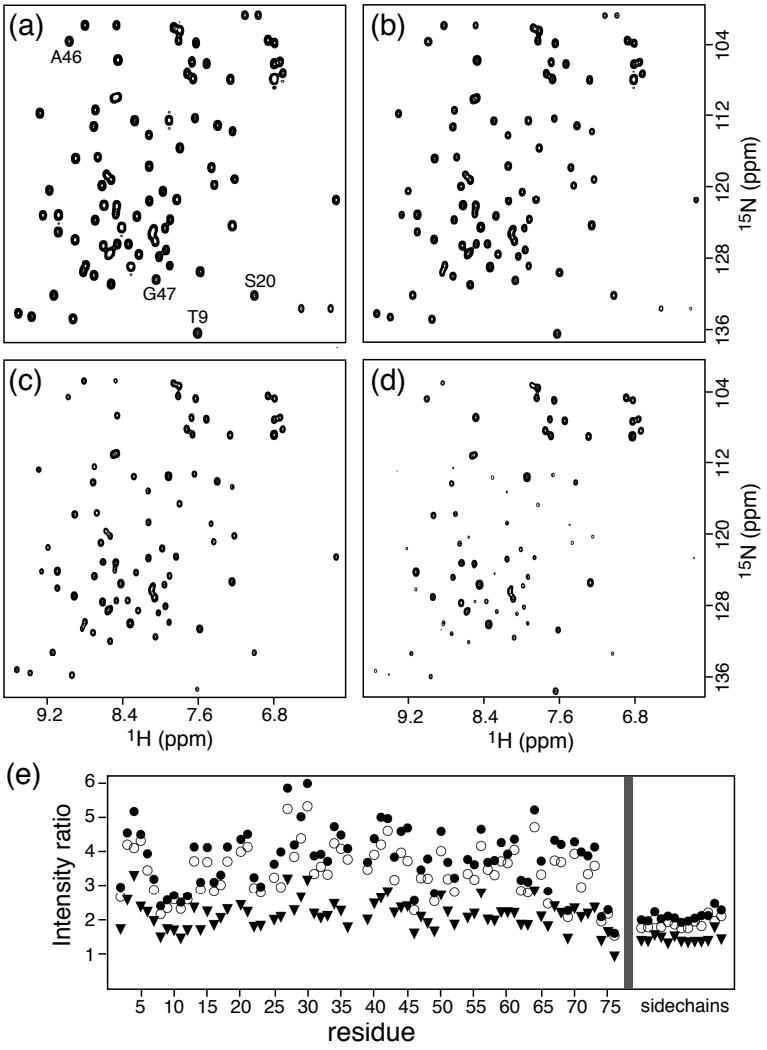


Figure 4.3 Comparison of 2D ^{1}H-^{15}N correlation spectra of ubiquitin recorded at 600MHz using (a) the SOFAST-HMQC pulse sequence of fig. 4.1a (α = 120°), (b) the LHSQC sequence of Pervushin et al. (2002), (c) a longitudinal-relaxation optimized se-afb-HSQC sequence and (d) a standard se-wfb-HSQC sequence as provided by the Varian Bio-pack. For se-afb-HSQC, the water flip-back pulse in the standard se-wfb-HSQC sequence was replaced by a band-selective EBURP flip-back pulse covering the ^{1}H frequency range from -1 to 5 ppm (see figure 3.2a). All spectra were acquired with a scan time of T_{scan} = 165 ms, and the acquisition times were set to t_1^{max} =28 ms, and t_2^{max} = 40 ms. ^{15}N-decoupling during t_2 was realized using WURST-2 (Kupce and Wagner (1995)) at an average field strength of $\gamma B_1/2\pi$=550 Hz. Additional 2D spectra were recorded with a longer t_1 acquisition time of t_1^{max} =50 ms (not shown). In (e) intensity ratios are plotted as a function of the peptide sequence: SOFAST-HMQC over se-wfb HSQC for t_1^{max} = 28 ms (filled circles) and t_1^{max} = 50 ms (open circles), and LHSQC over se-wfb HSQC (filled triangles).

intensities measured in the different 2D spectra of figure 4.3 shows a sensitivity gain for SOFAST-HMQC with respect to se-wfb-HSQC varying between 2 and 6, whereas this gain is only about 1.5 to 3.5 for LHSQC, and even less for se-afb-HSQC despite its $\sqrt{2}$-sensitivity advantage provided by the sensitivity-enhanced quadrature detection scheme. This sensitivity advantage of SOFAST-HMQC with respect to se-wfb-HSQC is slightly reduced when using longer t_1 acquisition times, e. g. t_1^{max}=50 ms (open circles) instead of t_1^{max}=28 ms (filled circles), because of the shorter transverse relaxation times of ^{1}H-^{15}N multiple quantum (MQ) coherence in HMQC-type sequences with respect to ^{15}N single-quantum (SQ) coherence evolution in HSQC-type sequences (see e.g. Cavanagh et al. (1995), page 429). The variations in the measured intensity ratio along the protein backbone are indicative of differences in the local proton density at the individual amide sites and/or differences in local dynamics. This effect has already been shown in figure 2.21, and chapter 8 will come back to this issue.

The principal conclusions from these experimental results are the following: (i) If optimized acquisition parameters (scan time, flip angle) are used, and moderate t_1 acquisition times are acceptable, SOFAST-HMQC yields the most sensitive ^{1}H-^{15}N correlation spectra of folded proteins. (ii) In the context of very fast data acquisition, SOFAST-HMQC provides a much higher sensitivity than se-wfb HSQC using the same scan times, and a similar sensitivity as se-wfb HSQC recorded with optimized inter-scan delays. The intensity curves shown in Figure 4.2 also provide some guidelines for setting up SOFAST-HMQC experiments. The highest sensitivity is obtained for scan times between 200 and 300 ms using a flip angle of α=120°, independent of the molecular size or magnetic field strength, whereas for shorter scan times ($T_{scan} < 200$ ms) flip angles of α=130°-150° are advantageous.

4.3.2 SOFAST-HMQC using cryogenic probes

The use of a cryogenic probe is very attractive in the context of fast data acquisition as cryogenic probes provide the required high sensitivity to record protein correlation spectra in a short overall experimental time. Unfortunately, the rf power necessary for the standard SOFAST-HMQC pulse sequence (figure 4.1a) at high magnetic field using short inter-scan delays results in high duty cycles, and it may be of interest to have a rf power-reduced experiment. We therefore propose an alternative sequence where the X-decoupling is replaced by an IPAP filter. This IPAP-SOFAST-HMQC sequence, shown in figure 4.1b, has allowed us to record a ^{1}H-^{15}N correlation spectrum on a ^{15}N-labelled sample of ubiquitin (0.9 mM, 25°C, pH 6.2) at submillimolar concentration on a 800 MHz spectrometer equipped with a cryogenic probe. The spectrum recorded in an overall experimental time of 12 seconds

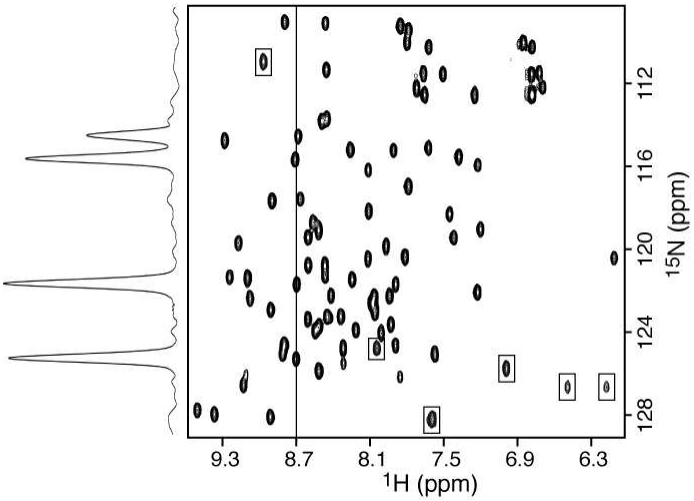


Figure 4.4 ^{1}H-^{15}N correlation spectrum of ^{15}N-labeled ubiquitin (0.9 mM, 25°C, pH 6.2) recorded on an 800 MHz INOVA spectrometer equipped with a cryogenic probe using the IPAP-SOFAST-HMQC sequence of Fig. 4.1b. The band-selective ^{1}H excitation (PC9) and refocusing (REBURP) pulses were centered at 8.0ppm covering a bandwidth of 4.0 ppm. The acquisition parameters were set to α=140°, Δ=5.4 ms, δ=1.2 ms, t_1^{max} = 22 ms, t_2^{max} = 40 ms, and t_{rec} = 1.0 ms. Forty complex data points were acquired in the t_1 dimension, n=80 + 4 dummy scans. Two data sets were recorded as explained in the caption to figure 4.1b, in an overall experimental time of 12 s. For data processing, the two raw data sets were first added and subtracted yielding the new data sets S1 and S2. Then a first order phase shift $\Delta\varphi_1 = +(2\pi \mathrm{sw} J_{HN})/2$ and $\Delta\varphi_2 = -(2\pi \mathrm{sw} J_{HN})/2$ was applied in the t_2 dimension to S1 and S2, respectively, with sw the spectral width in the ^{1}H dimension. Finally the two data sets were added and Fourier transformed as usual. Boxes indicate cross peaks with a ^{15}N frequency outside the chosen ^{15}N spectral width of 1800 Hz that are folded back into the spectrum.

is shown in figure 4.4. Data acquisition details are provided in the figure caption. The high S/N ratio obtained in this short experimental time is highlighted by the 1D trace extracted along the ^{15}N dimension, demonstrating the performance of the IPAP-SOFAST-HMQC for application on high-field NMR spectrometers equipped with cryogenic probes.

Although this rf power-reduced version may be safely applied to any cryogenically cooled probe, it suffers from two-fold longer minimal experimental times and a sensitivity loss of $\sqrt{2}$ with respect to the standard, ^{15}N decoupled version, as described above. In an attempt to circumvent these problems, we investigated in more detail the effects of fast-pulsing ^{15}N-decoupled SOFAST-HMQC on cryogenic probes. On the assayed probes (Varian ColdProbe) we find that even when using very short recovery delays, ^{15}N decoupled SOFAST-HMQC can be used without any restriction. Therefore, it is in general preferable to use the decoupled version

rather than the IPAP version.

However, some precautions are found necessary: The temperature of the probe initially rises when running fast-pulsing decoupled experiments. This may change the peak intensities. The internal feedback regulation of the probe controling system accounts for this rf heating, and after about one minute this down-regulation re-establishes the normal temperature (25 K) of the cold probe. However, the intensities extracted from experiments recorded during this time may vary, and in cases where accurate intensities are measured (e.g. for measuring kinetics) we therefore like to perform "dummy scans" (i.e. repetitions of the experiment without signal acquisition) for ca. one minute to reach stable conditions of the probe, before actually starting the measurement. Once stable conditions are reached we do not find problems concerning instability of the probe (t_1 noise).

Another important finding is the considerable detuning of the ^{15}N channel, which increases the 90° pulse length by about 15-20%. This seems to be a common feature on the current cryo-probes in experiments where high rf load is applied to the ^{15}N transmitter, and is also found in other experiments applying high B_1 fields to the heteronuclear transmitter channel (e.g. CPMG-type sequences used for measurements of ^{15}N transverse relaxation). We account for this detuning by calibrating the 90° pulse under fast-pulsing conditions.

4.3.3 Fast measurement of ^{1}H-^{15}N coupling constants

Another interest of the IPAP-SOFAST-HMQC sequence is that it allows fast measurement of H-X (^{1}H-^{15}N or ^{1}H-^{13}C) spin coupling constants from the ^{1}H frequency difference of cross peaks detected in the two sub-spectra corresponding to the α and β spin states of the attached hetero-nuclei. The quantification of one-bond spin coupling constants is especially interesting in the presence of an alignment medium that induces a dipolar contribution to the line splitting (Tjandra and Bax (1997)). These residual dipolar couplings provide valuable information about molecular structure and dynamics (Blackledge (2005)).

For fully-protonated, partially-aligned protein samples, one-bond (scalar and residual dipolar) spin coupling constants can be accurately measured from the line splitting observed along the heteronuclear frequency dimension in spin-state-selective H-X correlation experiments. In IPAP-SOFAST-HMQC the couplings are measured in the directly detected ^{1}H dimension, which is intrinsically less accurate. The ^{1}H line shape for aligned protein samples is generally asymmetric because of the presence of numerous ^{1}H-^{1}H dipolar interactions inducing a fine structure in the peak shape, and cross-correlated relaxation effects responsible for unequal line widths and intensities of the individual multiplet lines (Brutscher

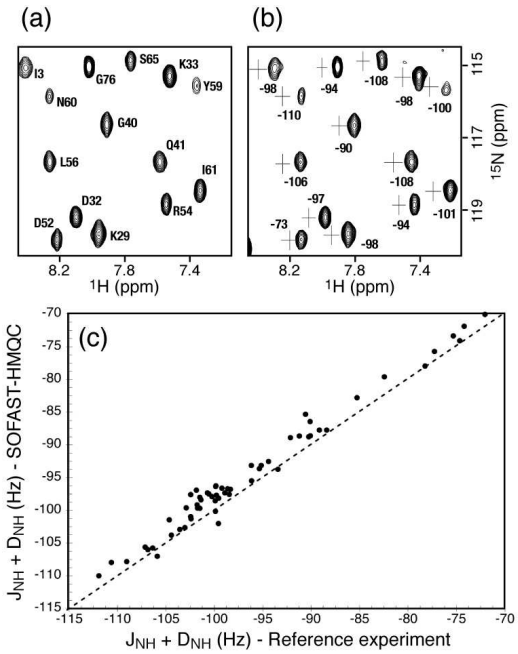


Figure 4.5 Experimental demonstration of fast ^{1}H-^{15}N coupling measurement using SOFAST-HMQC. Spectra were recorded at a 800 MHz spectrometer equipped with a cryogenic probe on a 2 mM sample of ubiquitin aligned in an alcohol mixture (Ruckert and Otting, 2000) at 30°C. Two data sets were acquired using the IPAP-version of SOFAST-HMQC as explained in the caption to figure 4.1b. A small part of the two sub-spectra containing the downfield and upfield doublet components are shown in (a) and (b), respectively. A cross in (b) indicates the position of the corresponding doublet peak in spectrum (a). In addition, residue number and type information is given in (a), and the measured coupling constants (in Hz) are provided in (b). The same acquisition parameters were used as reported for the spectrum of figure 4.4 yielding an overall experimental time of 12 seconds. In (c) the measured coupling constants are plotted against the line splittings measured in the ^{15}N dimension from a pair of TROSY-type ^{1}H-^{15}N spin-state selective correlation spectra (Weigelt, 1998) recorded in an experimental time of 25 minutes.

(2000)). This induces some systematic error in the coupling measurements. In order to evaluate experimentally the performance of IPAP-SOFAST-HMQC, in terms of accuracy and precision for the measurement of ^{1}H-^{15}N (scalar and residual dipolar) coupling constants, we have performed an IPAP-SOFAST-HMQC experiment on a partially-aligned 2mM sample of ^{15}N-labeled ubiquitin, dissolved in aqueous solution containing a 5% C12E5/hexanol mixture (r=0.85) (Ruckert and Otting (2000)). The quadrupolar ^{2}H line splitting observed for these conditions was about 30 Hz at 30°C. The IPAP-SOFAST-HMQC sub-spectra, corresponding to the down- and upfield ^{1}H doublet components, are shown in Figs. 4.5a and b. A total of 58 coupling constants could be quantified from non-overlapping cross peaks in these spectra, obtained in an overall experimental time of only 12 seconds. These results can be compared to couplings measured along the ^{15}N dimension of TROSY-type spin-state selective ^{1}H-^{15}N correlation spectra (Weigelt (1998)) recorded in an experimental time of 25 minutes. The correlation of the two data sets, shown in Fig. 4.5c, is quite good, although a systematic error (as mentioned above) tends to slightly underestimate the couplings measured using the fast IPAP-SOFAST-HMQC approach. In conclusion, IPAP-SOFAST-HMQC is certainly not the best method for precise and accurate measurement of ^{1}H-^{15}N residual dipolar couplings for NMR structure determination, but the quality of the data, nevertheless, demonstrates the potential of SOFAST-HMQC for probing structural changes in proteins that occur on a time scale of seconds via the fast measurement of one-bond spin couplings in partially aligned protein samples.

4.3.4 ^{1}H-^{13}C SOFAST-HMQC of methyl groups

The SOFAST-HMQC experiment is not limited to ^{1}H-^{15}N correlation spectroscopy, but can also be used to record ^{1}H-^{13}C correlation spectra of sub-ensembles of aliphatic and aromatic protons in the protein. As an attractive first example we demonstrate here its application to methyl groups. Methyl groups are valuable probes of structure and dynamics as they are dispersed throughout the primary sequence and often located in the hydrophobic core of proteins. Due to the rapid rotation, methyl groups possess three equivalent protons, which intrinsically yields a threefold intensity improvement. Their location in the hydrophobic core should make them excellent probes for following folding reactions, where the hydrophobic core is formed starting from a highly solvent-exposed structural ensemble. It has also been shown that methyl groups are useful probes for the study of molecular interfaces, and drug binding (Hajduk et al. (2000)). Thus there is an interest in methods that allow recording very fast ^{1}H-^{13}C methyl correlation spectra.

The relatively good spectral separation of methyl ^{1}H resonances (-0.5 to 1.5 ppm) from other aliphatic proton resonances makes the extension of SOFAST-

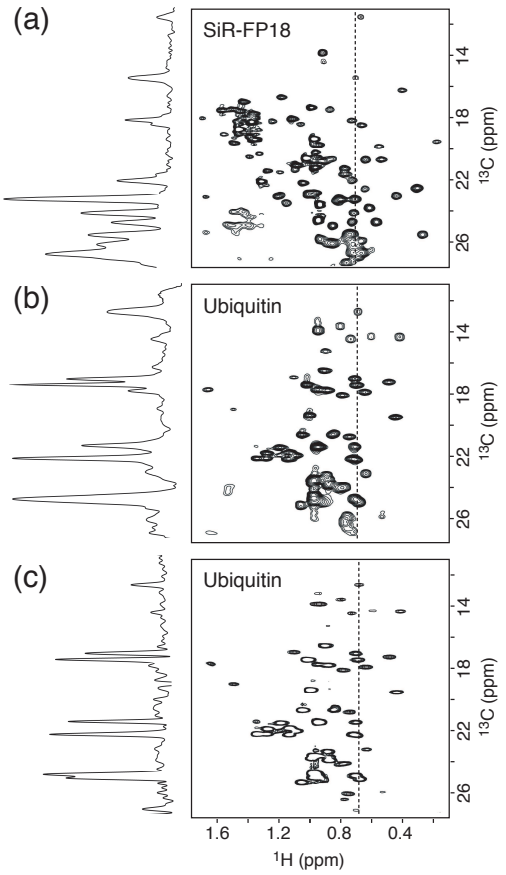


Figure 4.6 Methyl ^{1}H-^{13}C SOFAST-HMQC spectra of (a) SiR-FP18 (1.5 mM, 30°C, 800 MHz), (b) ubiquitin (2 mM, 25°C, 600 MHz), both recorded using the pulse sequence of figure 4.1a, and (c) ubiquitin (2 mM, 25°C, 600 MHz) using the CT-SOFAST-HMQC sequence of Fig. 4.1c. The band-selective ^{1}H excitation (PC9) and refocusing (r-SNOB) pulses were centered at 0.7ppm covering a bandwidth of 2.0 ppm. The acquisition parameters at 800MHz were set to α=150°, Δ=3.8 ms, δ=2.5 ms, t_1^{max} = 30 ms, t_2^{max} = 40 ms, and t_{rec} = 1.0 ms. 120 complex data points were acquired in t_1, n=240 + 4 dummy scans, yielding a total experimental time of 16s for the spectrum in (a). At 600 MHz the following parameters were used: α=150°, Δ=3.8 ms, δ=3.4 ms, t_1^{max} = 30 ms (23 ms), t_2^{max} = 40 ms, and t_{rec} = 1.0 ms. 80 (60) complex data points were acquired in t_1, n=160 (120) + 4 dummy scans, yielding a total experimental time of 10s for the spectra in (b) and (c). The values in parenthesis correspond to the settings of the CT experiment shown in (c). Multiple-band selective ^{13}C decoupling during methyl ^{13}C frequency editing was realized by a train of CA-WURST-2 (Kupce and Freeman (1996)) pulses of lengths τ_p=5 ms, centered at 37, 53, and 70 ppm, and covering bandwidths of 12 ppm each (Van Melckebeke et al. (2004)).

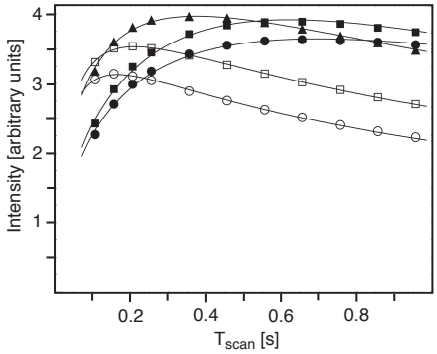


Figure 4.7 Sensitivity plotted as a function of the scan time (T_{scan}) obtained with different ^{1}H-^{13}C methyl correlation experiments for ubiquitin (2 mM, 25°C) at 600 MHz. 2D spectra were recorded using the SOFAST-HMQC sequence of fig. 4.1a with flip angles of α = 90° (filled squares), 120° (triangles), 140° (open squares) and 150° (open circles), and using a standard HSQC sequence (filled circles). Acquisition parameters are identical to those used for the spectrum in Fig. 4.6b except that no homonuclear ^{13}C decoupling was applied during t_1. The plotted data points correspond to the sum of the intensities measured for 27 well-resolved methyl cross peaks in the 2D spectra

HMQC to this spin system straightforward. Figures 4.6a and b show ^{1}H-^{13}C methyl correlation spectra of SiR-FP18 (1.5mM, 30°C, 800MHz), and of ubiquitin (2 mM, 25°C, 600 MHz), respectively, recorded with the pulse sequence of figure 4.1a. The spectra were recorded on standard triple-resonance (non cryogenic) probes in an overall experimental time of 16s (Fig. 4.6a) and 10s (Fig.4.6b). Contrary to amide ^{1}H-^{15}N SOFAST-HMQC, for methyl spectra the use of an r-SNOB pulse shape for ^{1}H refocusing yielded better results than the use of a REBURP pulse. This finding already indicates that the longitudinal relaxation enhancement effect is less pronounced for methyl protons than for amide protons. Further experimental details are given in the figure caption. Methyl correlation spectra of proteins often suffer from severe signal overlap. To increase spectral resolution, additional homonuclear ^{13}C decoupling (Y channel in figure 4.1a) was applied during t_1 as described in detail previously (Van Melckebeke et al. (2004)). This homo-decoupling removes J_{CC} line splittings in the ^{13}C dimension except for methyls at the δ position of isoleucine and leucine residues and therefore also yields a gain in sensitivity. An alternative to homonuclear decoupling is the use of CT ^{13}C editing, allowing increased spectral resolution in the ^{13}C dimension at the expense of sensitivity. Because of the relatively long transverse relaxation times of ^{1}H-^{13}C MQ coherences in methyl groups, CT editing only slightly reduces the sensitivity of the experiment for small proteins such as ubiquitin. A ^{1}H-^{13}C CT-SOFAST-HMQC recorded using the pulse sequence of figure 4.1(c) in an experimental time of 10 s is shown in figure 4.6c. 1D traces extracted along the ^{13}C dimension highlight the sensiti-

vity and spectral resolution obtained by the SOFAST-HMQC experiments. Methyl proton spin-lattice relaxation is dominated by the dipolar interaction among the methyl ^{1}H and ^{13}C that is time modulated by the fast methyl rotation, and only to a smaller extent by interactions with non-methyl protons in the surrounding. Therefore the longitudinal ^{1}H relaxation enhancement using selective methyl ^{1}H manipulation in SOFAST-HMQC is minor, yielding only a slight signal enhancement with respect to a standard HSQC experiment. This is demonstrated in figure 4.7 showing the average sensitivity of the methyl cross peaks as a function of the scan time. Much larger effects are expected for other more rigid aliphatic or aromatic ^{1}H sites in the protein. Still, for short inter-scan delays the optimized flip angle in SOFAST-HMQC provides on average a 50% signal increase with respect to standard HSQC.

SOFAST-HMQC may also prove useful for application to large molecular systems with a high level of deuteration and specific ^{1}H and ^{13}C labels at the methyl positions. (Rosen et al. (1996)), where the MQ coherence evolution yields TROSY-type line narrowing (Tugarinov et al. (2003)), and the optimized flip angle provides increased sensitivity for short inter-scan delays.

4.4 Conclusions

SOFAST-HMQC provides a robust new tool for biomolecular NMR. It allows to adjust the acquisition time of two-dimensional heteronuclear correlation spectra of proteins as a function of the available sensitivity down to a minimal time of a few seconds. It also provides higher signal to noise ratio for a given experimental time than other ^{1}H-^{15}N correlation experiments. Different implementations of SOFAST-HMQC have been presented, that are optimized for the use on standard or cryogenic NMR probes, for additional homonuclear decoupling in the indirect frequency dimension, and for the measurement of (scalar and residual dipolar) spin coupling constants. In view of the continuously increasing sensitivity of biomolecular NMR instruments, it can be expected that SOFAST-HMQC will become a widespread tool for high-throughput and real-time NMR investigations of protein structure, dynamics, and kinetics.

5 Speeding Up Three-Dimensional Protein NMR Experiments to a Few Minutes

5.1 Band-selective Excitation Short Transient HNCO/CA

5.1.1 Introduction

Triple-resonance experiments, that correlate backbone amide ^{1}H, ^{15}N to CA, CO, CB atoms, are a prerequisite for resonance assignment of proteins. Here, the longitudinal relaxation optimized approach used in the 2D SOFAST-HMQC experiment is extended to triple resonance experiments by incorporating additional coherence transfers to carbon spins. A different approach than in SOFAST-HMQC is used, to create in single-quantum ^{15}N coherence, rather than multiple-quantum ^{1}H-^{15}N coherence, which is more prone to transverse relaxation losses (see e.g. Cavanagh et al. (1995), page 429).

It is demonstrated that the sensitivity provided by these experiments is sufficient to record 3D H-N-CO and 3D H-N-CA spectra of ^{13}C/^{15}N-labeled proteins within a few minutes of data collection.

5.1.2 BEST HNCO/CA experiments

The new Band-selective Excitation Short-Transient (BEST) HNCO/CA pulse sequences are shown in Figure 5.1. The coherence transfer pathway is identical to standard sensitivity-enhanced HNCO/CA, and it makes use exclusively of band-selective pulses to the amide proton spins or pairs of broadband inversion pulses. The choice of band-selective pulses has been optimized in terms of minimal perturbation of spins resonating outside the chosen frequency band, and minimal signal

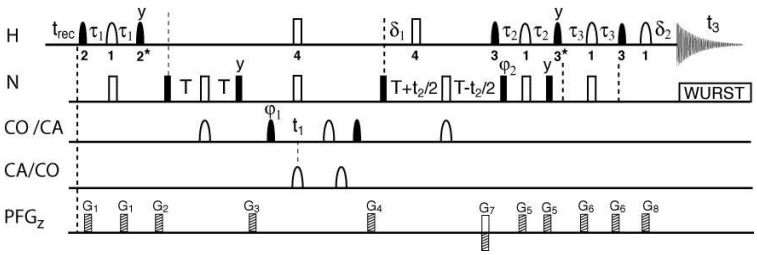


Figure 5.1 BEST-HNCO/CA pulse sequences to record HNCO or HNCA correlation spectra of proteins. Filled and open pulse symbols indicate 90° and 180° rf pulses. All selective ^{1}H pulses are centered at ∼8.0ppm covering a band width of 4.0ppm with the following shapes: (1) REBURP (Geen and Freeman (1991)) (2) PC9 (Kupce and Freeman (1994)) and (3) EBURP-2 (Geen and Freeman (1991)). A star indicates a flip back pulse obtained by time and phase inversion. The particular pulse shapes have been chosen to minimize signal loss because of transverse spin relaxation and B_1-field inhomogeneities. Open squares on ^{1}H indicate BIP-720-50-20 pulses (Smith et al. (2001)). CO pulses have the shape of the center lobe of a sinx/x function, whereas CA pulses are applied with a rectangular shape and zero-excitation at the CO frequency. The relative durations of G_7 and G_8 are given by the gyromagnetic ratios $G_7/G_8=\gamma_H/\gamma_N$. Quadrature detection in t_1 and t_2 is obtained by time-proportional phase incrementation φ_1=x,y and φ_2=x,-x, respectively. In addition, the sign of G_7 is inverted for experiments with φ_2=-x. At 600 MHz the following pulse lengths and delays were used: EBURP-2: 1.9 ms, RE-BURP: 2.0 ms, and PC9: 3.0 ms, τ_1 =100 μs, τ_2 =1.4 ms, τ_3=400 μs, T =12-15 ms, and δ_1 =$t_2/2+\delta$, with the delay δ ∼ 200 μs adjusted to suppress J_{NH} coupling evolution during the N→C transfer delay 2T. The maximal evolution times are set to t_1^{max} ∼ 10-15 ms, t_2^{max} ∼ 20-25 ms, and t_3^{max} ∼ 40 ms. Adiabatic ^{15}N WURST-40 (Kupce and Wagner (1995)) decoupling ($\gamma B_1/2\pi \sim 500$ Hz) is applied during detection.

losses due to relaxation and pulse imperfections. This ensures the high sensitivity that is needed for short repetition delays t_{rec}.

Choice of band-selective pulses

The pulse sequence in figure 5.1 has the same coherence transfer pathway as standard HNCO/CA experiments, based on an initial INEPT transfer block, and a concluding planar-mixing sequence element, and figure 5.2 illustrates the properties of the chosen selective pulses during the relevant sequence elements.

The initial ^{1}H-^{15}N INEPT-transfer step requires an excitation, a refocusing and a flip-back pulse (figure 5.2). We have chosen a REBURP shape for refocusing because of its clean off-resonance performance with respect to other pulse shapes, e.g. r-SNOB. In order to minimize ^{1}H transverse relaxation delays, we have chosen PC9 pulses for excitation and flip back. Unlike pure-phase pulses that yield uniform phase over the pulse bandwidth, like EBURP, E-SNOB..., these pulses allow chemical shift and J_{NH}-coupling evolution during part of the pulse duration

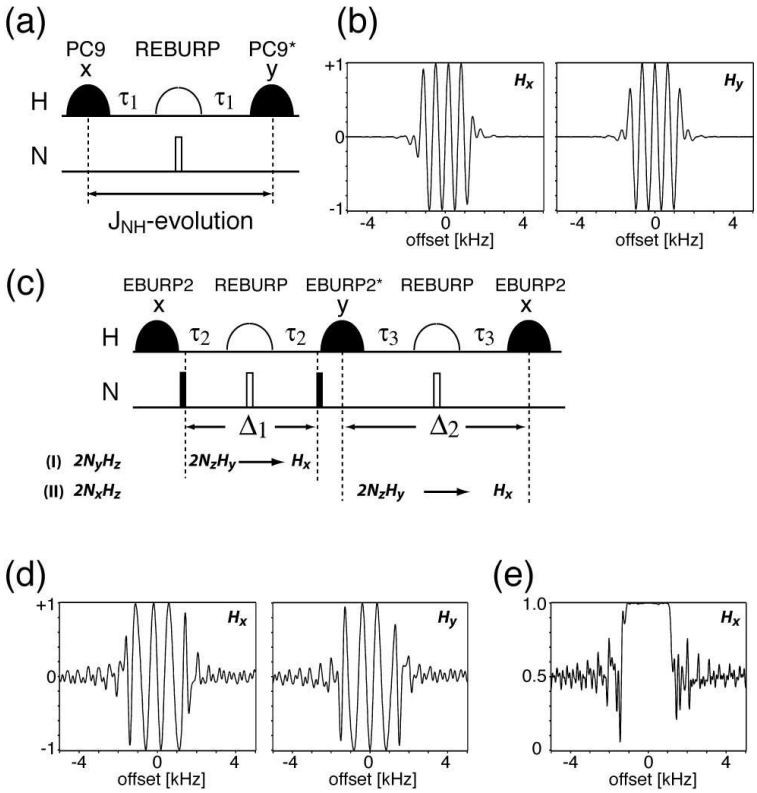


Figure 5.2 Selection of optimal pulse shapes for BEST-HNCO/CA. (a) and (b): Optimization of the initial INEPT transfer. (a) Coupling evolution is active during the delay indicated by an arrow, including the duration of the REBURP pulse and half of each PC9 pulse. (b) Bloch simulations of the PC9 excitation profile: H_x and H_y components are plotted as a function of frequency offset. The PC9 pulse length was set to 3.0 ms, corresponding to an excitation band width of 4.0 ppm at 600 MHz. From these simulations, one can estimate the phase gradient that is accumulated during the pulse, and thus the fraction of the pulse duration, during which chemical shift evolution is active. In addition, experimental optimization of the sensitivity as a function of τ_1 reveals that J_{NH} coupling is also active during half of the pulse duration. (c) Band-selective planar mixing transfer sequence required for sensitivity-enhanced ^{15}N quadrature detection. The 2 relevant coherence transfer pathways corresponding to ^{1}H-^{15}N back-transfer of the orthogonal ^{15}N coherences are represented on the bottom. (d) and (e) Bloch simulations of the planar mixing sequence. ^{1}H transverse coherence created (d) by the first pulse sandwich EBURP2 REBURP EBURP2*, and present (e) after the complete planar mixing sequence block shown in (c) simulated for coherence transfer pathway II. Again, chemical shift and J_{NH}-coupling evolution is active during ∼50% of the EBURP2 pulse length, that allows to shorten the transfer delays τ_3 accordingly. The resulting planar-mixing sequence thus makes optimal use of the long duration of the band-selective ^{1}H pulses (∼2.0 ms at 600 MHz), and avoids extensive signal loss due to B_1-field inhomogeneities.

(∼50% of the pulse length, as estimated from simulations, figure 5.2(b), and experiment). The coupling evolution is therefore active during 2 times ca. 1.5 ms (at 600 MHz) during the PC9 pulses, and during the whole duration of the REBURP pulse. This allows to reduce the coupling evolution delay τ_1 between the pulses, as compared to a sequence using pure-phase pulses, and reduce transverse relaxation losses during this INEPT transfer step.

The final planar mixing sequence (Kay et al. (1992)) retains two orthogonal projections for detection, which increases the sensitivity by a factor of up to $\sqrt{2}$. The two relavant coherence transfer pathways are indicated in figure 5.2(c).

Again REBURP pulses have been chosen for refocusing. For the 90° rotations, EBURP-2 pulse shapes are used that create pure-phase coherence from z magnetization. The first EBURP-2 pulse is applied at the end of the N→C transfer delay (2T). J_{NH}-coupling evolution then occurs between the two 90° ^{15}N pulses ($\Delta_1 \sim 1/(2J_{NH})$). In principle, general rotation pulses are required for the two remaining 90° pulses. The central 90° pulse (applied along the y-axis) has to restore the H_x coherence along the z-axis (pathway I) without affecting the $2N_yH_y$ coherence that needs to be transferred during the delay Δ_2 ∼$1/(2J_{NH})$ (pathway II). The last 90° pulse has to create H_y coherence from H_z (pathway I) without affecting H_x (pathway II).

In principle, general rotation pulses, e.g. UBURP, Q5, can be used to achieve this goal, but these pulses have been shown to be particularly sensitive to B_1-field inhomogeneities and pulse miscalibration. We thus explored possibilities to use pulse shapes that are more robust, although not optimized as general rotation pulses. The use of EBURP-2 pulses proved successful; although this pulse shape is optimized for excitation ($H_z \rightarrow H_x$) and its time-reversed counterpart achieves the opposite action ($H_x \rightarrow H_z$), it also have the following unexpected and unreported properties:

(i) H_y coherence present at the beginning of a (flip back) EBURP-2* pulse (phase set to y) evolves as if the pulse was a spin-lock field (during ∼ half of the pulse length) followed by a free evolution (and J_{NH} coupling) delay. This property is used for the central 90° pulse in the scheme of figure 5.2(c). The resulting superposition of H_x and H_y of magnetization from coherence pathway II (which is present as H_y before this pulse) is shown in figure 5.2(d).

(ii) Similarly, and related to (i), the superposition of H_x and H_y, created by this pulse (and a subsequent π pulse) is left in the transverse plane during a EBURP-2 pulse (phase x). Again, it appears that this pulse can be thought of as a free

evolution delay followed by a spin-lock field. These two properties allow to create a pure-phase H_x magnetization from $2H_zN_x$ (pathway II) with the sequence element EBURP-2 (x) - REBURP - EBURP-2* (y) - REBURP - EBURP-2 (x) (figure 5.2(e)). While doing so, the coherence pathway I follows the normal properties of the EBURP-2 pulses: flip-back of a pure-phase H_x coherence by the central EBURP-2*, and excitation $H_z \rightarrow H_y$ by the concluding EBURP-2.

(iii) Another interesting feature is that a *flip-back* pulse EBURP-2* applied to H_z creates transverse magnetization with a close to linear phase gradient over the excitation profile. In other words, the flip-back pulse applied to H_z can be thought of as a shorter pulse followed by a free evolution delay (similar to the PC9 pulse). This property allows to use part of the duration of the pulse for coupling and chemical shift evolution. This is exploited in a related selective experiment for the sensitive measurement of ^{1}H-^{1}H coupling constants (chapter 6).

The two π pulses used for ^{1}H decoupling (marked with "4" in figure 5.1) do not necessarily have to be applied in a band-selective manner. When applied as hard pulses, they result in a 360° rotation of the aliphatic proton spins, which still allows a net flip-back of aliphatic magnetization. Non-selective pulses have advantages due to their shorter durations. To investigate whether hard pulses or selective pulses at these two positions perform better, we implemented the two ^{1}H decoupling pulses as selective pulses (REBURP, ISNOB), hard pulses and optimized Broadband Inversion Pulses (BIP, Smith et al. (2001)). We find the best results when using BIP pulses:[1] they afford a much cleaner inversion profile as compared to rectangular hard pulses, resulting in significant sensitivity gains in the fast-pulsing regime (see figure 5.3).

This has been noted before, and it can generally be recommended to replace hard inversion pulses by BIP. In addition, it has also been shown that replacing refocusing pulses by two BIP pulses yields signal gains on the order of a few percent, and they are therefore attractive in many experiments (Diercks et al. (2005); Keppetipola et al. (2006)).

5.1.3 Enhanced sensitivity afforded by BEST-HNCO/CA

The sensitivity of BEST-HNCO/CA is plotted in Figure 5.4 as a function of the recycle delay and compared to results obtained for a standard pulse scheme us-

[1] A BIP pulse termed BIP-720-50-20 denotes a pulse that has a duration equivalent to a 720° square pulse, compensating for 50% frequency offset and 20 % B_1 field inhomogeneity.

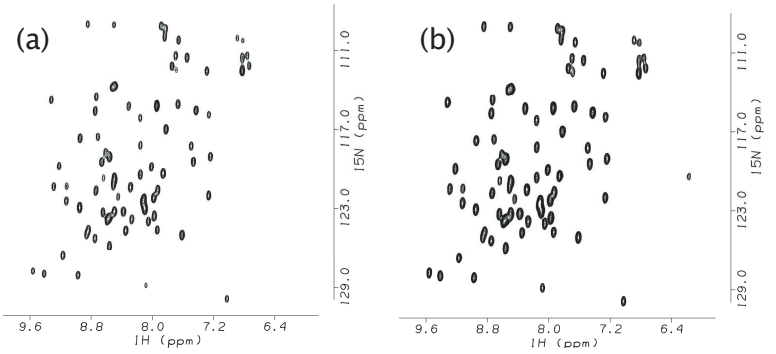


Figure 5.3 Comparison of ^{1}H-^{15}N projections of HNCO spectra of ubiquitin recorded with the BEST-HNCO experiment using (a) rectangular 180° pulses and (b) BIP 720-50-20 pulses for proton decoupling. The data were acquired in the fast pulsing regime (t_{rec} set to 100 ms).

ing two different ^{13}C/^{15}N-labeled protein samples: ubiquitin and SiR-FP18 (*E. coli* sulfite reductase flavodoxine-like domain, PDB: 1YKG). Signal gains of up to a factor of 2.8 are observed for both proteins when using the BEST sequences with high repetition rates. The absolute maximum of these curves is comparable for the standard and the BEST experiments. Therefore, BEST experiments are especially interesting in cases where fast data acquisition is required.

As deuteration is now widely used when studying larger proteins, we have also evaluated the performance of BEST-HNCO/CA for a deuterated (ca. 75%) sample of SiR-FP18. Despite the small number of remaining protons, a significant sensitivity gain is observed, as shown in Figure 5.5. These data encourage the use of BEST-HNCO/CA when studying partially deuterated protein samples.

5.1.4 Application to proteins: Acquisition of HNCO/CA spectra in a few minutes

To demonstrate that the sensitivity provided by the BEST sequences is sufficient to record 3D HNCO and HNCA spectra in a few minutes of acquisition time on standard high-field NMR spectrometers, we have performed experiments on small- to medium-sized ^{13}C/^{15}N-labeled proteins at millimolar concentration.

Figure 5.6(a) shows N-CO planes extracted from a 3D H-N-CO correlation spectrum recorded for ^{13}C/^{15}N ubiquitin on a 600 MHz spectrometer equipped with a room-temperature triple-resonance probe. The data set was acquired with 1 scan per (t_1,t_2) increment in an experimental time of only 10 min. Note that the BEST se-

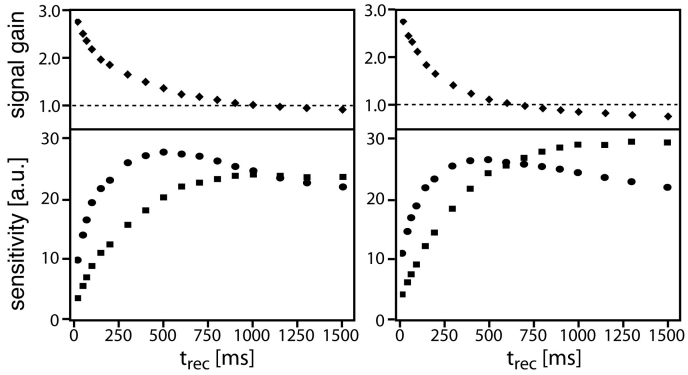


Figure 5.4 Sensitivity (i.e. signal-to-noise per unit time) measured using different recycle delays t_{rec} for BEST-HNCO/CA (circles), and standard wfb-se-HNCO/CA (squares) as implemented in the Varian BioPack. The intensity ratios (BEST over standard) are plotted on top. 2D ^{1}H-^{15}N data sets were recorded at 600 MHz for (left curves) ubiquitin (8.6 kDa, pH 6.2, 25°C) and (right curves) SiR-FP18 (18 kDa, pH 7.5, 30°C). Each point corresponds to the sum of all amide cross peak intensities.

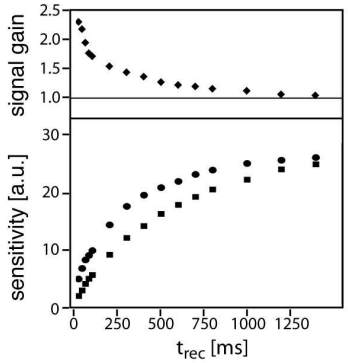


Figure 5.5 Sensitivity plotted as a function of the recycle delay t_{rec} for BEST-HNCO/CA (filled circles) and standard se-wfb HNCO/CA (filled squares) for deuterated (∼75%) SiR-FP18 (18kDa, pH 7, 25°C). The sensitivity ratio of the two experiments is plotted on top. The data were obtained as an intensity sum of all cross peaks in the spectrum. Spectra were recorded on a Varian DirectDrive 600MHz spectrometer equipped with a room-temperature probe.

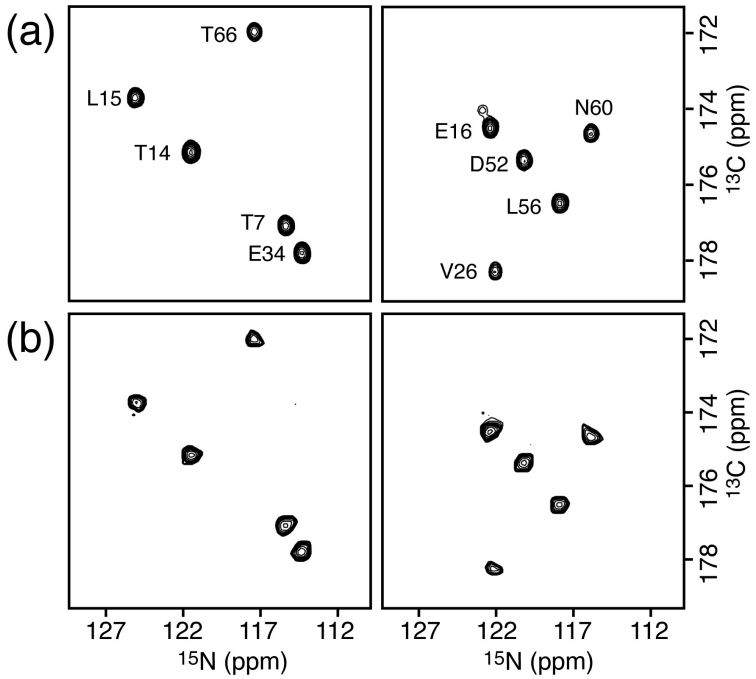


Figure 5.6 N-CO planes extracted from (a) a 3D BEST-HNCO spectrum recorded in 10 min, and (b) a BEST-HNCO spectrum reconstructed using a minimal value algorithm (Kupce and Freeman (2004a)) from a set of 6 2D projections, recorded in an overall time of only 2 min. For spectrum (a) 30 (^{13}C) $\times$ 27 (^{15}N) complex points were acquired with t_{rec} = 50 ms. 2D projections were obtained by setting $\Delta t_1 = |cos(\alpha)| / SW_{tilt}$ and $\Delta t_2 = |sin(\alpha)| / SW_{tilt}$, with α the projection angle and SW_{tilt} the spectral width in the projected dimension. All data were recorded at 600 MHz, 25°C on the 1.9 mM ubiquitin sample.

quences yield good water suppression within a single scan. In order to obtain spectra of comparable S/N ratio using a standard sequence, either a ∼5-times longer recycle delay or a ∼10-times longer acquisition time would have been required (see Figure 5.4(a)). All expected cross peaks were observed in this BEST-HNCO spectrum with an average S/N ratio of 50:1.

In addition, we have recorded a set of 6 2D projections for angles α=0°, ±30°, ±60°, and 90° in an acquisition time of 20 seconds per projection. A first possibility of extracting the correlated frequency triplets from these projections is to reconstruct the 3D spectral space using retro-projection techniques (Kupce and Freeman (2004a)). Figure 5.6(b) shows the same N-CO planes as plotted in (a) extracted from the reconstructed 3D H-N-CO spectrum, obtained in a total acquisition time of only 120 seconds.

An alternative strategy consists in extracting the frequency information directly

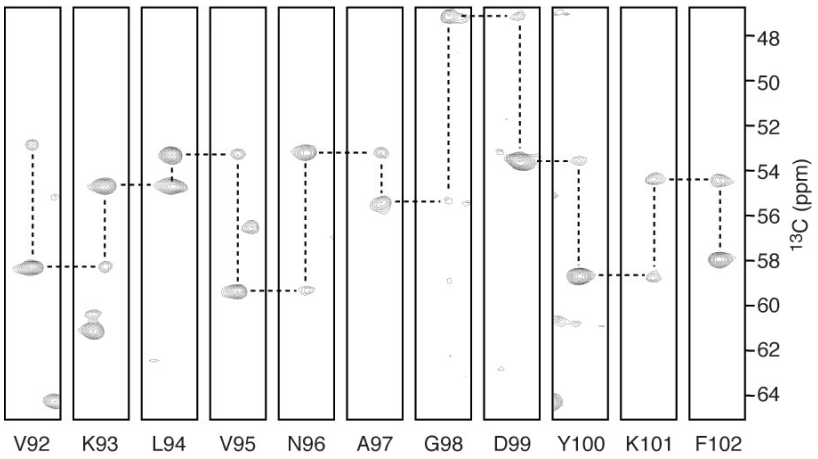


Figure 5.7 HN-CA strips extracted from a 3D BEST-HNCA spectrum of 1.8 mM $^{13}C/^{15}N$-labelled SiR-FP18 (30°C). 30 (13C) × 27 (15N) complex points were acquired with t_{rec} = 150 ms in an experimental time of 15 min.

from a set of projection spectra using symmetry criteria as implemented in several assignment programs (Hiller et al. (2005a); Malmodin and Billeter (2005b); Morelle et al. (1995)). For small globular proteins, 2 or 3 projection spectra are generally sufficient to retrieve the complete 3D spectral information (Bersch et al. (2003)). This approach further reduces the required acquisition time to less than one minute.

To further investigate the applicability of BEST-HNCO/CA to larger proteins in the 100-200 residues range, we recorded a 3D HNCA spectrum of SiR-FP18 (167 residues). The experiment was performed on a 1.8 mM $^{13}C/^{15}N$-labelled sample, in an experimental time of 15 min. Figure 5.7 shows a series of strips extracted at the H, N frequencies of residues V92 to F102 illustrating the potential of sequential resonance assignment using fast 3D BEST-HNCA.

5.1.5 Conclusions

In summary, we have shown that 3D H-N-CO and H-N-CA protein correlation spectra can be recorded within a few minutes, sometimes even less, of data collection. We have also shown that the BEST experiments are compatible with projection NMR techniques to further reduce acquisition times. It is equally possible to combine BEST-HNCO/CA with other fast acquisition techniques such as Hadamard-type or spatially encoded ^{15}N and/ or ^{13}C frequency labeling.

The BEST sequences can be easily extended to other correlation experiments involving only amide protons, e.g. sequential, intra-residue, and bi-directional

H-N-CA, H-N-CO, and H-N-CB correlation experiments required for backbone resonance assignment, as well as scalar and residual dipolar coupling, and auto- and cross-correlated relaxation rate constant measurements.

5.2 Extension of the BEST concept to a full set of triple-resonance experiments for protein resonance assignment

5.2.1 Introduction

For full backbone resonance assignment in proteins, a set of H-N-C correlation experiments has to be collected. Here, the BEST-concept introduced in section 5.1 is extended to the complete series of sequential, intra-residue, and bi-directional H-N-C (C = CO, C^{α}, or C^{β}) correlation experiments. We demonstrate that 3D data sets can be recorded in about 15 to 40 min using uniform time domain sampling. These acquisition times proved to be sufficient to detect all expected correlation peaks in the spectra of a 2 mM sample of ubiquitin recorded on a 600 MHz spectrometer equipped with a cryogenically cooled probe. Interestingly, these BEST experiments do not require any particular data processing tools.

5.2.2 BEST triple resonance experiments

Figure 5.8 shows BEST-HNCO, BEST-HN(CO)CA, and BEST-HN(CO)CACB pulse sequences for sequential correlation of the amide group of residue i with the CO, C^{α}, or C^{β} carbon of residue (i-1).

Figure 5.9 shows BEST-iHN(CA)CO, BEST-iHNCA, and BEST-iHNCACB pulse sequences for correlation of the amide group with the CO, C^{α}, or C^{β} of the same residue. The pulse sequence elements used to perform intra-residue N→C^{α} coherence transfer have been introduced and described previously (Brutscher (2002); Nietlispach (2004); Tossavainen and Permi (2004)).

Finally, figure 5.10 shows BEST-HNCA, BEST-HNCACB, and BEST-HN(CA)CO pulse sequences used to correlate the amide group with the CO, C^{α}, or C^{β} carbons of both the same and the preceding residue. As usual, the parentheses in the experiment name indicate a nuclear spin that is involved in the coherence transfer pathway, but not frequency labeled. Note that the CACB-type experiments can be tuned to yield correlation peaks either with only the C^{β} ((CA)CB-type, $\Delta_2=1/(4J_{CC})$), or with both the C^{α} and C^{β} carbons (CACB-type, $\Delta_2=1/(8J_{CC})$). The coherence transfer pathways of these BEST experiments are of the out and back type, and identical to the corresponding hard-pulse based correlation experiments that are routinely used for resonance assignment of $^{13}C/^{15}N$ labeled proteins.

The only difference, besides the use of band-selective ^{1}H pulses, is the absence of any composite ^{1}H decoupling during ^{15}N and $^{13}C^{\alpha,\beta}$ transverse coherence evolution times that is generally applied to reduce relaxation-induced signal losses

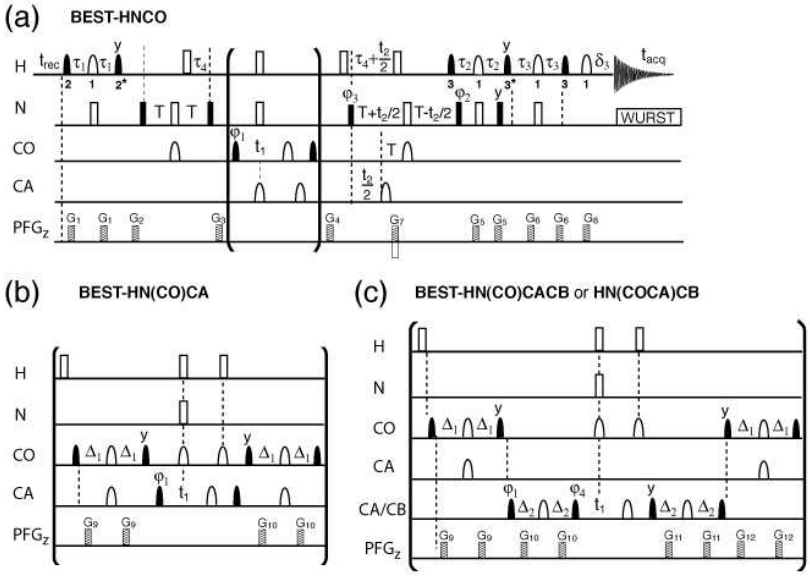


Figure 5.8 BEST pulse sequences to record sequential correlation spectra of proteins: (a) BEST-HNCO, (b) BEST-HN(CO)CA, and (c) BEST-HN(CO)CACB and BEST-HN(COCA)CB. The four experiments differ by the sequence elements represented inside brackets. Filled and open pulse symbols indicate 90° and 180° rf pulses. Unless indicated, all pulses are applied with phase x. All selective ^{1}H pulses are centered at 8.2 ppm, covering a bandwidth of 4.0 ppm, with the following shapes: (1) REBURP, (2) PC9, and (3) EBURP-2. A star indicates a flip back pulse obtained by time and phase inversion. Open squares on ^{1}H indicate BIP-720-50-20 pulses. CO pulses have the shape of the center lobe of a sinx/x function, whereas CA pulses are applied with a rectangular shape and zero excitation at the CO frequency. The delays are set to: τ_1=2.4 ms 0.5 δ_1, τ_2=2.4 ms, τ_3=2.4 ms 0.5 δ_2, τ_4=2.7 ms, T=14.5 ms, Δ_1=4.5 ms, Δ_2= 3.5 ms for CACB-type or Δ_2=7.1 ms for (CA)CB-type experiments. The delays δ_1, δ_2, and δ_3, correspond to the pulse lengths of the PC9, E-BURP2, and G8 gradient, respectively. These settings do not take into account the durations of the 180° rf pulses applied during the transfer delays that need to be subtracted. Pulsed field gradients, G1G8 are applied along the z-axis (PFGz) with durations of 200 μs to 2 ms and field strengths ranging from 540 G/cm. The phase cycling is: φ_1= x,-x, φ_2= x, φ_3= 2(x), 2(-x), φ_4= y, and the receiver φ_{rec}= -x, x, x, -x. In the 3D applications shown here, only a 2-step cycle of φ_1 and φ_{rec} is performed. The relative durations of G7 and G8 are given by the gyromagnetic ratios G7/G8=γ_H/γ_N. Quadrature detection in t_1 is obtained by time-proportional phase incrementation of φ_1 and φ_4 (if present) according to TPPI-States. For quadrature detection in t_2, echo-antiecho data are recorded by inverting the sign of gradient G7 and phaseφ_2. In addition, phase φ_3 is inverted for every second t_2 increment.

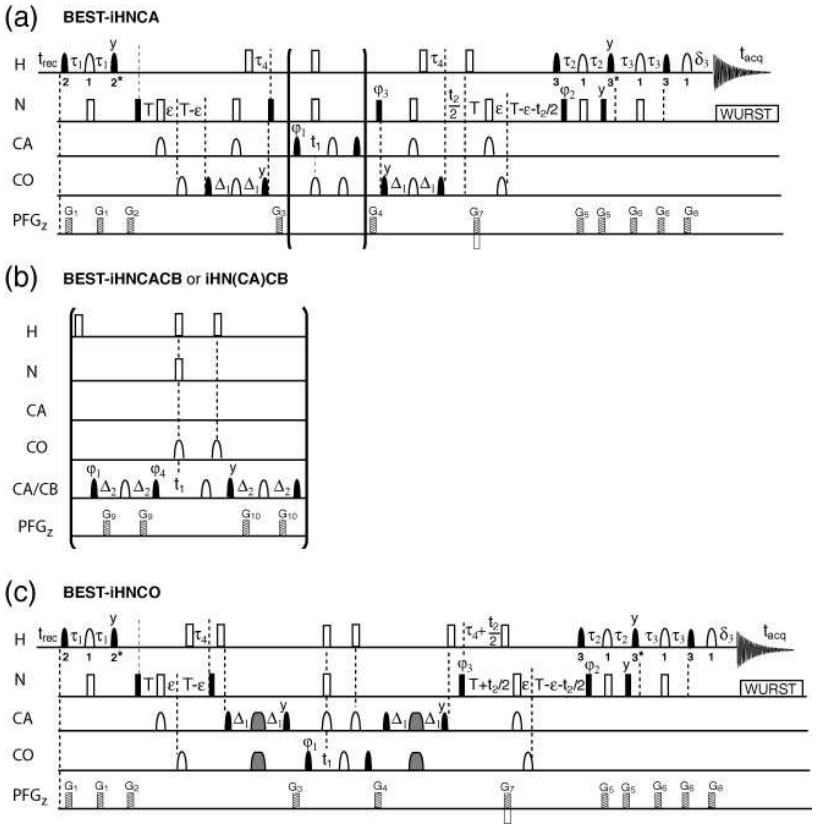


Figure 5.9 BEST pulse sequences to record intra-residue correlation spectra of proteins: (a) BEST-iHNCA, (b) BEST-iHNCACB and BEST-iHN(CA)CB, and (c) BEST-iHNCO. The BEST-iHNCACB and BEST-iHN(CA)CB experiments differ from BEST-iHNCA by the sequence elements represented inside brackets. The delays and phases are identical to those given in the caption of figure 5.8, except for T=17-20 ms (for sequences a and b) and T=22-25 ms (for sequence c), and ε=T-17 ms. The grey shaped pulses applied in sequence (c) on the CA and CO channels to achieve selective CA→CO transfer are applied with a REBURP profile covering a band width of 20 ppm, and centered at 175 ppm and 56 ppm, respectively.

(Tossavainen and Permi (2004)). In BEST experiments, the signal loss arising from the perturbation of the aliphatic spin polarization by the composite ^{1}H decoupling would counterbalance the signal gain due to the longer transverse relaxation times of N_x (C_x) with respect to $2N_xH_z$ ($2C_xH_z$) coherence. In a recent work, Diercks et al. (2005) have compared the performance of different broadband ^{1}H decoupling sequences with respect to the perturbation of aliphatic ^{1}H polarization, and obtained the best results with an XY-16 sequence. We have tested the performance of our BEST experiments using either BIP-based XY-16, band-selective IBURP-based XY-8, or no ^{1}H composite decoupling during the ^{15}N-^{13}C transfer periods. For ubiquitin using short inter-scan delays ($t_{rec} < 500$ ms), highest sensitivity was obtained for the sequences without additional ^{1}H decoupling (data not shown), and therefore ^{1}H decoupling was omitted in the final BEST experiments.

The initial coherence transfer steps for the sequential, intra-residue, and bi-directional BEST experiments are given by the following equations:

$$H_z^i \overset{J_{NH}}{\rightarrow} 2H_z^iN_z^i \overset{J_{NC'},J_{NH}}{\longrightarrow} 2N_z^iCO_z^{i-1} \overset{J_{NC'},J_{NH}}{\longrightarrow} 4N_z^iCO_z^{i-1}CA_z^{i-1} \overset{\cdots}{\rightarrow} \ldots \quad (5.1)$$

$$H_z^i \overset{J_{NH}}{\rightarrow} 2H_z^iN_z^i \overset{J_{NC'},J_{NC^\alpha},J_{C'C^\alpha},J_{NH}}{\longrightarrow} 4N_z^iCO_z^{i-1}CA_z^i \overset{\cdots}{\rightarrow} \ldots \quad (5.2)$$

$$H_z^i \overset{J_{NH}}{\rightarrow} 2H_z^iN_z^i \overset{J_{NC^\alpha},J_{NH}}{\longrightarrow} 2N_z^iCA_z^{i-1} + 2N_z^iCA_z^i \overset{\cdots}{\rightarrow} \ldots \quad (5.3)$$

A special case is the BEST-iHNCO experiment of figure 5.9c for which the coherence transfer pathway is given by:

$$H_z^i \overset{J_{NH}}{\rightarrow} 2H_z^iN_z^i \overset{J_{NC'},J_{NC^\alpha},J_{NH}}{\longrightarrow} 8N_z^iCO_z^{i-1}CA_z^{i-1}CA_z^i \overset{J_{C'C^\alpha}}{\rightarrow} 8N_z^iCA_z^{i-1}CA_z^iCO_z^i \quad (5.4)$$

5.2.3 Sensitivity of BEST experiments

In order to compare the performance of the BEST sequences presented here with standard hard pulse sequences, we have focused on the intra-residue iHN(CA)CB experiment which presents the worst-case scenario for the BEST approach because of the long transverse N→CA ($\sim$ 90 ms) and CA→CB ($\sim$ 30 ms) transfer delays involved, and the absence of ^{1}H decoupling during these periods. Series of 1D spectra were acquired as a function of the scan time (pulse sequence length plus recycle delay) for both BEST (figure 5.9b) and standard pulse sequences on a 2 mM sample of ^{13}C/^{15}N-labeled ubiquitin (pH 6.3, 25°C) on a 600 MHz spectrometer equipped with a cryogenic triple-resonance probe. For the standard pulse sequence, band-selective HN pulses in the sequence of figure 5.9b were replaced by hard pulses, and additional composite ^{1}H decoupling, and water flip-back pulses

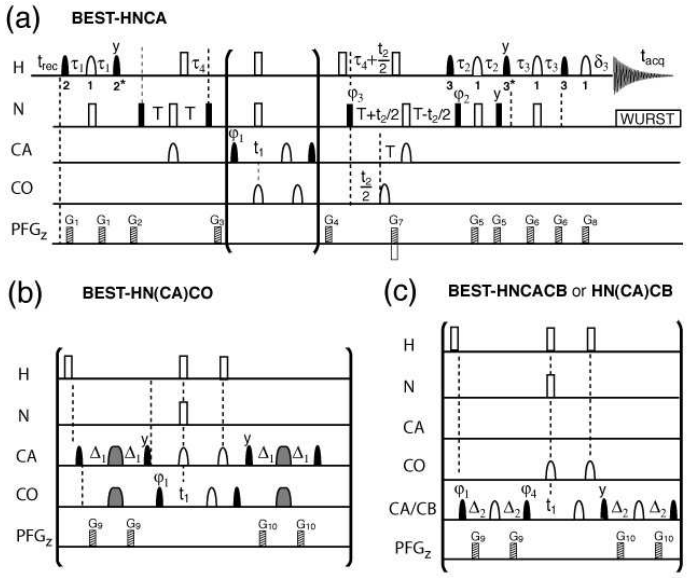


Figure 5.10 BEST pulse sequences to record bi-directional correlation spectra of proteins: (a) BEST-HNCA, (b) BEST-HN(CA)CO, and (c) BEST-HNCACB and BEST-HN(CA)CB. The four experiments differ by the sequence elements represented inside brackets. The delays and phases are identical to those given in the caption of figure 5.8, except for T=12 ms. The grey shaped pulses applied in sequence (b) on the CA and CO channels to achieve selective CACO transfer are applied with a REBURP profile covering a band width of 20 ppm, and centered at 175 ppm and 56 ppm, respectively.

were added as usual. The CA→CB transfer delays were tuned to Δ_2=$1/4J_{CC}$ for complete transfer to the Cβ carbons. The 1D spectra were integrated in the range 7.2 to 9.2 ppm. The measured intensities normalized for equal acquisition times are plotted as a function of the scan time (T_{scan}) in figure 5.11a. As expected from the longitudinal relaxation enhancement effect, the maximum of these sensitivity curves is shifted from T_{scan} ≈1.5 s for the standard pulse sequence to T_{scan} ≈0.5 s for the new BEST sequence. This results in a small absolute sensitivity gain (∼25%) for the BEST experiment if optimal inter-scan delays are used for both sequences (optimal sensitivity regime). Higher sensitivity gains are achieved for high repetition rates of both pulse sequences. Average sensitivity gains of ∼50% and ∼100% are observed for scan times of 500 ms and 350 ms, respectively, while keeping a high overall sensitivity for the BEST experiment. Even higher gains are obtained for scan times $<$ 350 ms (fast pulsing regime) at the expense of a reduced overall sensitivity. In order to evaluate the sensitivity gain for individual amide sites along the polypeptide chain we have also recorded 2D BEST-iHN(CA)CB and standard iHN(CA)CB data sets for an inter-scan delay of t_{rec}=200 ms (T_{scan}=370 ms). The spectra are shown in figures 5.11c and 5.11d, and a histogram of the measured in-

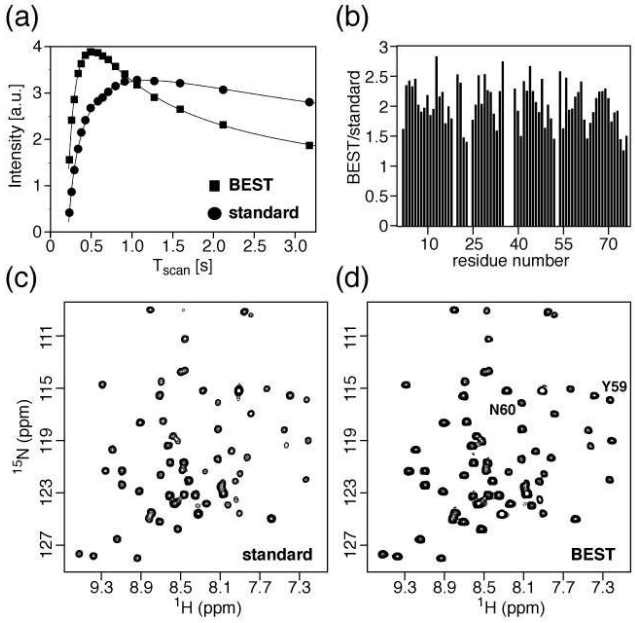


Figure 5.11 (a) Sensitivity plotted as a function of the scan time T_{scan} for iHN(CA)CB spectra recorded using the BEST implementation (squares) and a standard hard-pulse version (circles). 1D spectra were recorded on ubiquitin (pH 6.3, 25°C) at 600 MHz using a cryoprobe. Integrals in the range 7.2 to 9.5 ppm are displayed. 2D ^{1}H-^{15}N spectra recorded using (c) standard and (d) BEST versions of the iHN(CA)CB experiments were obtained with a recycle delay trec set to 200 ms (T_{scan}=370 ms). The total duration of both experiments was identical (∼17min). The intensity ratios obtained for individual correlation peaks are plotted in panel (b).

tensity ratios (BEST/standard) is plotted in figure 5.11b. These data confirm the average gain of about a factor of 2 as predicted by the sensitivity curves of figures 5.11a. The sensitivity gain varies from a factor of 1.5 up to a factor of ∼3.0 for individual amides, clearly demonstrating the interest of the BEST approach when high repetition rates are desired to reduce the overall experimental time required to record a complete set of 3D H-N-C correlation spectra. Even higher gains are expected for the other experiments of the BEST series.

5.2.4 Application to the fast collection of assignment experiments

A set of 11 3D BEST H-N-C spectra has been acquired on a 2 mM sample of ^{13}C/^{15}N-labeled ubiquitin (pH 6.3, 25°C) on a 600 MHz spectrometer equipped with a cryogenic triple-resonance probe. The inter-scan delay was again set to t_{rec}=200 ms for all experiments, presenting a good compromise between high sensitivity and high repetition rates. In addition, we have used ASCOM optimization

Table 5.1 Acquisition parameters and statistics of 3D H-N-C spectra recorded on a 2 mM sample of ^{13}C-^{15}N labeled ubiquitin on a 600 MHz spectrometer equipped with a cryogenically cooled triple-resonance probe

3D BEST experiment	SW_N/SW_C (kHz)	No. of complex points n_N/n_C	Exp. time (min)	No. of peaks observed/expected[a]	Relative S/N ratio
HNCO	0.772/1.2	19/25	21	69/69	100[d]
HN(CO)CA	0.772/2.0	19/30	27	69/69	37
HNCA	0.772/2.0	15/30	20	138/138	22/9
					intra/seq
HN(COCA)CB	0.772/8.0	15/50	38	65/65	20
HN(CO)CACB	0.772/8.0	15/50	36	134/134	24/14
					C^α/C^β
iHNCA	0.772/2.0	19/30	28	69/69	15
iHN(CA)CO	0.772/1.2	15/25	25	65/69[b]	9
iHN(CA)CB	0.772/8.0	15/50	39	64/64	12
iHNCACB	0.772/8.0	15/50	38	133/133	13/9
					C^α/C^β
HN(CA)CO	0.772/1.2	19/25	17	130/138[b]	11/5
					intra/seq
HNCACB	0.772/8.0	15/50	34	263/267[c]	19/11-8/5
					intra C^α/C^β - seq C^α/C^β

The relative signal to noise (S/N) ratios were calculated for the average peak intensity in each experiment and normalized with respect to the HNCO experiment. The numbers of peaks expected in each spectra were calculated on the basis of residues with significant intensities in the ^{15}N-HSQC spectra.

[a] Among the 76 residues the 3 prolines and the N-terminal residues are not detected. Residues 24 and 53 have been excluded from the analysis because of extensive line broadening, as well as residue 36 for which the amide ^{1}H frequency is on the edge (6.15 ppm) of the excitation bandwidth chosen for the band-selective ^{1}H pulses.

[b] Cross-peaks with C^α of glycine residues have low intensity or are missing.

[c] Four sequential correlation peaks are missing (HN_9-CA_8, HN_9-CB_8, HN_{25}-CA_{24}, HN_{65}-CA_{64}).

[d] The absolute average signal-to-noise ratio measured for correlation peaks in the BEST-HNCO spectrum was 800:1.

(Lescop et al. (2007), see page 219) to minimize the ^{15}N spectral width without creating any additional peak overlap in the ^{1}H-^{15}N correlation spectrum. Using a 2-step phase cycle for axial peak suppression this resulted in acquisition times ranging from 17 to 39 min per 3D experiment depending on the spectral width and resolution chosen for the ^{13}C dimension. The acquisition parameters, record-

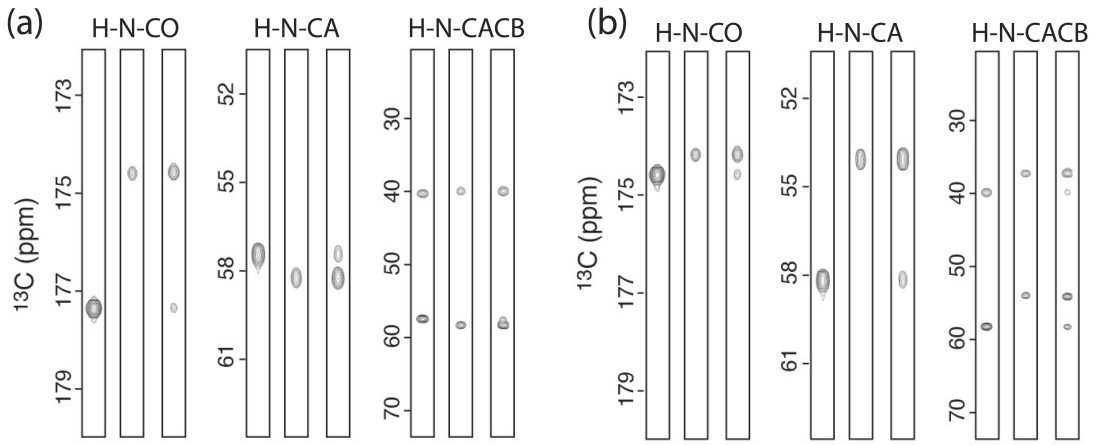


Figure 5.12 $^{1}H^{N}$-^{13}C strips extracted at the $^{1}H^{N}$/^{15}N chemical shifts of residues Y59 (panel a) and N60 (panel b) from nine 3D BEST experiments used for backbone chemical shift assignment. From the left to the right: HNCO, iHNCO, HN(CA)CO, HN(CO)CA, iHNCA, HNCA, HN(CO)CACB, iHNCACB and HNCACB. The spectra were recorded as indicated in table 1. Positive and negative correlation peaks are shown in black and grey, respectively.

ing times, number of peaks, and the relative signal to noise ratios observed in these spectra are summarized in table 5.1. The major conclusions are that all expected correlation peaks were observed in these spectra, except for 4 missing sequential correlation peaks in the HNCACB spectra. The undetected cross peaks in the BEST-iHN(CA)CO and BEST-HN(CA)CO spectra involve glycine residues for which the C chemical shift evolution is not refocused by the band-selective 180° pulses applied during the CA→CO transfer steps. Note that the bandwidth and carrier frequency of the amide ^{1}H pulses need to be adjusted to cover the complete chemical shift range of a given protein. Amide ^{1}H chemical shifts at the edge or outside the chosen excitation bandwidth will result in cross peaks of reduced or no intensity for this residue. On the other hand, widening the excitation band may reduce the sensitivity of the BEST experiment because of partial perturbation of aliphatic and water protons.

In summary, we have presented a set of 3D BEST H-N-C experiments that yield significantly increased sensitivity for high repetition rates with respect to standard pulse sequences. We have demonstrated that these BEST pulse sequences allow recording a complete set of 3D correlation spectra required for sequential protein resonance assignment in only a few hours. Short overall acquisition times are of particular relevance for unstable protein samples degrading rapidly, and for monitoring fast processes. More generally, the BEST sequences allow adjusting the acquisition times to the intrinsic sensitivity of the experimental set up (sam-

ple and spectrometer). The economized spectrometer time can then be used advantageously for other less sensitive protein samples, or for spending more time recording NMR data providing quantitative information on the protein structure and dynamics (nOes, RDCs, relaxation rate constants, J coupling constants, etc). We have shown recently for the 18 kDa protein SiR-FP18 that the application of BEST sequences is not limited to small proteins, but it proves also advantageous in terms of sensitivity for larger proteins. Interestingly, even for a 75% deuterated sample of SiR-FP18, the longitudinal relaxation enhancement of BEST experiments still leads to sensitivity gains compared to standard experiments (figure 5.5).

5.3 Aliphatic flip-back performance of SOFAST HMQC and BEST experiments

The sensitivity gains of the proposed experiments arise mainly from longitudinal relaxation enhancement of amide proton spins. The efficiency of this enhancement critically depends on the fraction of magnetization of non-used (mainly aliphatic) spins that is retained in a thermal equilibrium state (see figure 2.20).

To analyze the flip-back performance of the proposed experiments we have acquired 1D spectra of ^{13}C bound protons immediately after the pulse sequence of SOFAST-HMQC or BEST-HNCO. Such spectra reveal the amount of aliphatic magnetization left at the conclusion of the pulse sequence (figure 5.13(a)). For comparison, we have also performed the same analysis for an experiment that uses hard pulses in combination with a number of band-selective flip-back pulses applied to the aliphatic and water ^{1}H spins. This experiment was found to have the best aliphatic flip-back performance out of a number of experiments proposed recently (Diercks et al. (2005)).

Figure 5.13(b) shows the fraction of polarization of ^{13}C-bound ^{1}H spins (H^C) left after H-N correlation experiments as a function of the recycle delay t_{rec} between scans. The value at very long t_{rec} reflects the intrinsic performance of the pulse sequences to recover H^C polarization. SOFAST-HMQC and BEST HNCO have an intrinsic flip-back performance of more than 95 and 80%, respectively. The lower value for BEST-HNCO can be assigned to the larger number of pulses, accumulating the effects of pulse imperfections.

At shorter recycle delays, the amount of recovered longitudinal H^C magnetization drops, meaning that the perturbed ^{1}H magnetization does not recover completely between subsequent scans, and reaches a steady-state after a few repetitions. Two reasons can be identified for the perturbation of H^C spins: (i) pulse imperfections, which lead to a less-than-complete flip-back in one scan (as seen from the values at long t_{rec}) accumulate in the fast pulsing regime, and (ii) the H^C polarization is reduced by cross-relaxation from the excited amide spins (see figure 2.14(c)). However, due to their good intrinsic flip-back performance, SOFAST-HMQC and BEST-HNCO/CA retain more than 70 and 60% of the longitudinal H^C magnetization in the fast pulsing regime, respectively.

The same analysis performed for the experiment using aliphatic flip-back pulses (Diercks et al. (2005)) reveals a lower level of retained aliphatic polarization in the fast-pulsing limit; finally a simple aliphatic flip-back HSQC sequence (shown in figure 3.2(a)) retains less than 10% of the aliphatic polarization under fast-pulsing

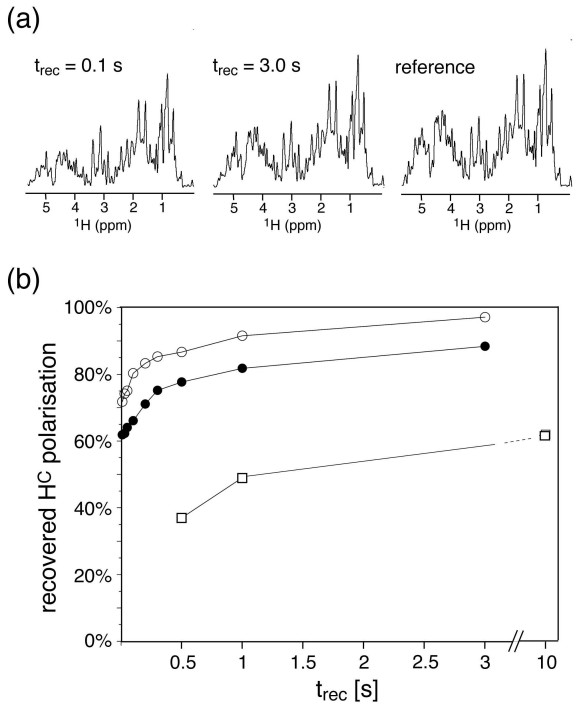


Figure 5.13 (a) 1D spectra of the aliphatic proton region of ubiquitin recorded with a ^{1}H-^{13}C HSQC sequence at 600 MHz. The spectra were acquired immediately after the last pulse of a BEST HNCO experiment performed for about half a minute with the respective recycle delay t_{rec} to reach steady state conditions. For the reference spectrum on the right, the BEST-HNCO block was omitted. (b) Fraction of the aliphatic proton polarization present at the end of a BEST-HNCO experiment (filled circles), SOFAST HMQC (open circles), and a optimized HNCO pulse sequence making extensive use of aliphatic flip-back pulses (Diercks et al. (2005), open squares). For recycle delays t_{rec} longer than 3 seconds the fraction of recovered H^C polarization does not change significantly and the shown values reflect the intrinsic capability of the pulse sequence to restore unused H^C polarization.

conditions (data not shown), which also explains that only moderate gains are achieved with this pulse sequence (see figure 4.3(c)).

In conclusion, minimal perturbation of unused aliphatic proton spins seems to be best achieved using carefully chosen selective pulses applied to the amide spins. In contrast, pulse sequences that make extensive use of hard pulses in combination with flip-back pulses perform less well.

6 Sensitivity-optimized experiment for the measurement of residual dipolar couplings between amide protons

6.1 Abstract

High signal to noise is a necessity for the quantification of NMR spectral parameters to be translated into accurate and precise restraints on protein structure and dynamics. An important source of long-range structural information is obtained from ^{1}H-^{1}H residual dipolar couplings measured for weakly aligned molecules. For sensitivity reasons, such measurements are generally performed on highly deuterated protein samples. Here we show that high sensitivity is also obtained for protonated protein samples if the pulse schemes are optimized in terms of longitudinal relaxation efficiency and J-mismatch compensated coherence transfer. The new sensitivity-optimized quantitative J-correlation experiment yields important signal gains reaching factors of 1.5 to 8 for individual correlation peaks when compared to previously proposed pulse schemes.

6.2 Introduction

Residual dipolar couplings (RDCs), Blackledge (2005); Prestegard et al. (2004); Tjandra and Bax (1997)) complement or replace more classical NMR data such as nuclear Overhauser effects (nOes) and scalar coupling constants (J). Whereas nOe-based distance restraints are affected by indirect spin interactions (spin diffusion), and J-coupling-derived torsion angles rely on empirical Karplus curves, RDCs directly reflect the interaction strength between two nuclear spins that can

be translated into structural restraints of high accuracy using a well-known analytical formula. RDCs therefore provide access to more precise definition of local and long-range molecular structure. It has also been shown that RDCs, measured in several alignment media, are particularly attractive for RDC-based de novo protein fold determination (Beraud et al. (2002); Hus et al. (2001); Kontaxis et al. (2005)), and the characterization of molecular dynamics at an atomic level occurring at time scales of pico- to milliseconds (Bouvignies et al. (2005); Lakomek et al. (2005); Meiler et al. (2001)).

Nowadays RDCs are routinely measured between covalently attached nuclei, e.g. N-H, C-H, N-C, and C-C, mainly in the protein backbone, as part of the standard data set recorded for NMR structure elucidation. The distance and orientational information provided by long-range RDCs between more distant nuclei is still less exploited. It has been shown recently for proteins and nucleic acids that ^{1}H-^{1}H RDCs can be accurately measured between protons separated by more than 7 Å, provided that the ^{1}H spin coupling network in these molecules is simplified by means of either deuteration, band-selective homonuclear decoupling, or both (Boisbouvier et al. (2003); Meiler et al. (2001); Wu and Bax (2002)). Here, we will focus on the measurement of long-range RDCs between amide protons in small to medium sized proteins. The sensitivity of the experimental schemes proposed so far greatly benefits from perdeuteration followed by back-protonation of the labile (amide) hydrogen sites. Most NMR experiments, commonly used for the study of small to medium-sized proteins, do not require perdeuteration. In order to make amide ^{1}H-^{1}H RDC measurements more attractive for fully protonated (or only partially deuterated) samples we have developed a new pulse sequence that provides greatly improved sensitivity over existing methods. Signal enhancements, ranging from a factor of 1.5 to 8, are observed for individual correlation peaks, as illustrated for ^{15}N labeled ubiquitin under weak alignment conditions.

6.3 Sensitivity-enhanced double ^{15}N-edited ^{1}H-^{1}H correlation experiment

In contrast to experiments for the measurement of one- or two bond couplings, long-range ^{1}H-^{1}H RDCs are most conveniently measured by a quantitative J correlation experiment (Bax et al. (1994)), where part of the magnetization is transferred from proton H^A to proton H^B yielding a cross peak at the resonance frequency of H^B and a diagonal peak at the frequency of H^A in the final spectrum. The spin coupling constant is then easily retrieved from the ratio of the measured intensities

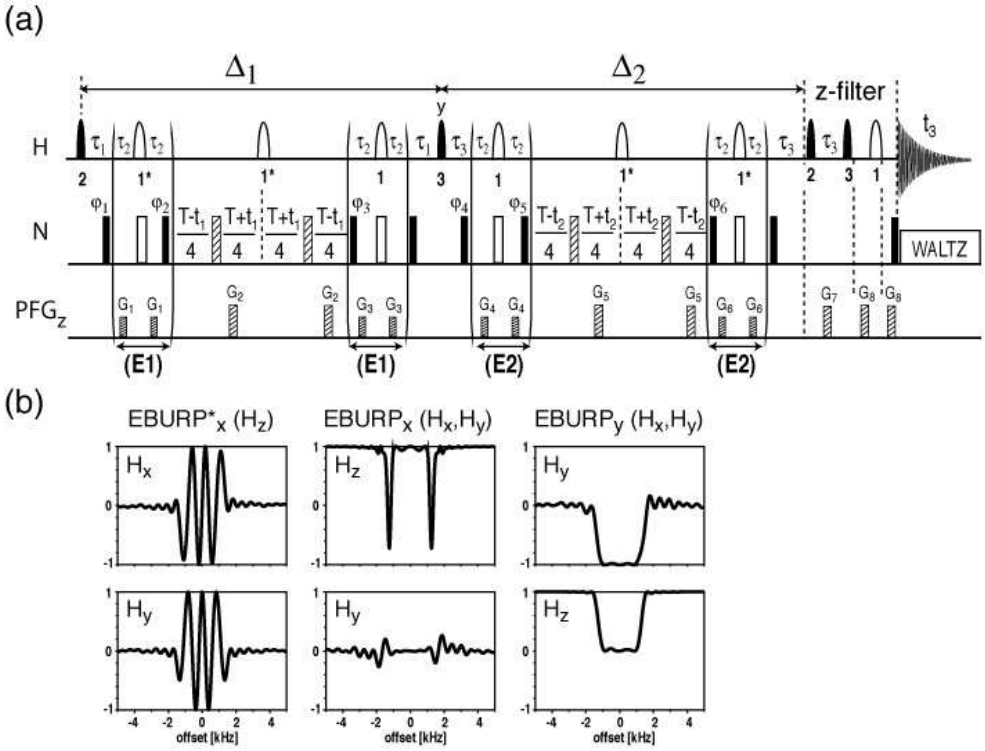


Figure 6.1 (a) Pulse sequence for the 3D BEST-Jcomp-HMQC2 experiment for the measurement of ^{1}H^{1}H residual dipolar couplings. The inserts labeled (E1) and (E2) correspond to J-mismatch compensation elements as explained in the text. The optional z-filter at the end suppresses $2H_xH_z$-type coherence prior to detection. All radio-frequency (rf) pulses are applied along the x-axis unless indicated. 90° and 180° rf pulses are represented by filled and open pulse symbols, respectively. The dashed ^{15}N pulses are broadband inversion pulses applied with a BIP-360-30-5 shape (Smith et al. (2001)). All shaped ^{1}H pulses are centered at 8.2 ppm covering a band width of 4.0 ppm corresponding to the amide ^{1}H spectral region. The following shapes are used: (1) REBURP, (2) time-reversed EBURP-2 (EBURP*), and (3) EBURP-2 (Geen and Freeman (1991)). For pairs of successive 180° rotations (1*) ISNOB5 pulses (Kupce et al. (1995)) instead of REBURP are applied because of their cleaner band-selective inversion profile. The transfer delays are adjusted to $\tau_1 = 1/(2J_{NH}) - \delta_1/2$, $\tau_2 = 1/(4J_{NH}) - \delta_2/2$, $\tau_3 = 1/(2J_{NH})$, $\tau_4 = 1/(2J_{NH}) - \delta_1$, and $T \approx 30 - 60$ ms, with δ_1 and δ_2 the lengths of the EBURP and REBURP pulses, respectively. The transfer delays Δ_1 and Δ_2 take into account ^{1}H-^{1}H coupling evolution during the selective pulses: $\Delta_1 = 2\tau_1 + 4\tau_2 + T$ and $\Delta_2 = 2\tau_3 + 4\tau_2 + T$. Pulsed field gradients, G1 - G8 are applied along the z-axis (PFGz) with a duration of 200 μs and field strengths ranging from 5 to 40 G/cm. Phase cycling: $\varphi_1 = x, -x, x, -x, \varphi_2 = y, -y, -y, y, \varphi_3 = 2y, 2(-y), \varphi_4 = x, \varphi_5 = 4y, 4(-y), \varphi_6 = 4y, 4(-y)$, and the receiver $\varphi_{rec} = x, -x$. Quadrature detection in the t_1 (t_2) dimension is obtained by simultaneous time-proportional phase incrementation of φ_1 (φ_4) and decrementation of φ_2 (φ_5) according to TPPI-STATES. (b) Bloch simulations illustrating the effect of EBURP-2 and time-reversed EBURP-2 pulses in the pulse sequence (a). Coupling evolution has not been taken into account for the simulations. The plotted offset profiles represent the ^{1}H spin state after different pulse combinations starting from pure z-magnetization (Hz): (left panel) EBURP*$_x$, (center) EBURP*$_x$-Δ-180°-Δ-EBURP$_x$, and (right) EBURP*$_x$-Δ-180°-ΔEBURP$_y$.

of cross- and diagonal peaks (see below). The new pulse sequence of Figure 6.1a is a modified version of the Semi-Selective HMQC2 (SS-HMQC2) experiment (Wu and Bax (2002), see figure 10.1 on page 217) that consists of two successive HMQC blocks sandwiching a COSY-type ^{1}H mixing pulse. In the following we will refer to our new experiment as BEST-Jcomp-HMQC2 (BEST: Band-selective-Excitation Short-Transient, Jcomp: J-mismatch compensated).

Before describing in more detail the new features of BEST-Jcomp-HMQC2, we will briefly summarize the relevant coherence transfer pathways common to the double-HMQC experiment. Amide ^{1}H chemical shift, scalar and dipolar coupling evolutions with aliphatic protons are refocused by the band-selective 180 ^{1}H pulses applied in the middle of each HMQC building block. In contrast, H^N-H^N RDCs evolve during the transfer delay Δ_1 as $H_y^A \rightarrow H_y^A \cos(\pi D_{HH}\Delta_1) + 2H_x^A H_z^B \sin(\pi D_{HH}\Delta_1)$, where D_{HH} is the residual dipolar coupling constant between spins H^A and H^B and interactions with other amide protons are neglected. The following 90° ^{1}H pulse then converts $2H_x^A H_z^B$ to $-2H_z^A H_x^B$ while leaving H_y^A (which gives rise to the diagonal peak) unaffected. During the subsequent Δ_2 transfer period, the antiphase coherence $-2H_z^A H_x^B$ partially refocuses to $H_y^B \sin(\pi D_{HH}\Delta_2)$ that is then detected as a crosspeak. Remaining $2H_z^A H_x^B$ and $2H_x^A H_z^B$ coherences that are apparent as dispersive antiphase signals in the ^{1}H (ω_3) dimension of the spectrum can be suppressed by an additional z-filter element that will be explained in more detail below. Contrary to the experiment of Wu and Bax (2002), where ^{1}H antiphase suppression is an integral part of the pulse scheme, removing the z-filter element from the sequence of Figure 6.1a does not alter the main coherence transfer pathways (see below). Therefore the experiment may also be performed without z-filter. During both ^{1}H-^{1}H dephasing times Δ_1 and Δ_2, the ^{15}N chemical shift of the covalently attached amide is labeled using HMQC-type ^{1}H-^{15}N correlation schemes. The residual dipolar coupling $D = D_{HH}$ is then obtained from the 4 peak intensities measured at positions (ω_1, ω_2, ω_3): $I^A_{cross}(\omega^A_N, \omega^B_N, \omega^B_H)$, $I^B_{cross}(\omega^B_N, \omega^A_N, \omega^A_H)$, $I^A_{diag}(\omega^A_N, \omega^A_N, \omega^A_H)$, $I^B_{diag}(\omega^B_N, \omega^B_N, \omega^B_H)$ using the following relation:

$$\frac{I^A_{cross} I^B_{cross}}{I^A_{diag} I^B_{diag}} = \frac{\lambda^2 \sin^2(\pi D\Delta_1)\sin^2(\pi D\Delta_2}{\lambda^2 \cos^2(\pi D\Delta_1)\cos^2(\pi D\Delta_2) + 2\lambda(1-\lambda)\cos(\pi D\Delta_1)\cos(\pi D\Delta_2) + (1-\lambda)^2} \tag{6.1}$$

with λ a correction factor taking into account the protonation level at the amide sites in the protein (Wu and Bax, 2002), e.g. λ = 0.9 for a 10%/90% D_2O/H_2O mixture.

Solving equation 6.1 for a measured intensity ratio yields the magnitude of the D_{HH} coupling constant, given by

$$|D_{HH}| = \frac{\gamma_H^2 \mu_0 h}{16\pi^3 r_{HH}^3} \left| A_a(3\cos^2\theta - 1) + \frac{3}{2} A_r \sin^2\theta \cos 2\varphi \right| \tag{6.2}$$

with A_a and A_r the axial and rhombic components of the alignment tensor, θ and φ the polar angles defining the H^N-H^N vector orientation with respect to the alignment frame, and r_{HH} the inter-proton distance.

The sensitivity of BEST-Jcomp-HMQC2 greatly benefits from two new features, longitudinal relaxation enhancement and J-mismatch compensation, that will be discussed in detail in the following.

6.3.1 Longitudinal relaxation enhancement

In slowly tumbling diamagnetic molecules such as proteins the efficiency of ^{1}H spin lattice relaxation is mainly determined by the dipolar-coupled ^{1}H spin network (^{1}H-^{1}H nOes). Therefore, if only a subset of protons is observed in a particular NMR experiment, spin-lattice relaxation can be enhanced by manipulating the remaining proton spins during the pulse sequence in a way that their spin state before detection is close to thermal equilibrium. The effective longitudinal spin-lattice relaxation of labile solvent-accessible amide ^{1}H is also influenced by chemical exchange with water hydrogens. Partial saturation of the water proton spins translates via this exchange mechanism to a reduced steady state polarization of the fast exchanging amides at the beginning of each scan. Again, longitudinal relaxation is enhanced when the water is in a relaxed state between subsequent scans. Shorter longitudinal relaxation times allow for higher repetition rates of the pulse sequence, and thus provide increased signal to noise (S/N) in a given amount of experimental time (higher sensitivity). Longitudinal relaxation enhancement has recently been successfully exploited for proteins in the context of resonance assignment experiments (Atreya and Szyperski (2004); Diercks et al. (2005); Pervushin et al. (2002); Schanda et al. (2006)), ultrafast two-dimensional data acquisition (sections 4, 5 and chapter 7), and fast characterization of structural compactness and heterogeneity of polypeptide chains (chapter 8).

In the pulse sequence of Figure 6.1a, only shaped ^{1}H pulses are applied that selectively manipulate amide ^{1}H while leaving aliphatic ^{1}H mostly unaffected. The BEST concept, consisting of the use of amide ^{1}H-selective pulses only, has been shown to yield superior results in terms of leaving aliphatic ^{1}H polarization unperturbed compared to hard-pulse-based sequences using additional selective "flip-back" pulses (section 5.3). In the sequence of Figure 6.1a, a single 180° ^{1}H pulse is applied with a REBURP shape, whereas for pairs of consecutive 180° ^{1}H pulses ISNOB5 pulse shapes are preferred because of their better off-resonance per-

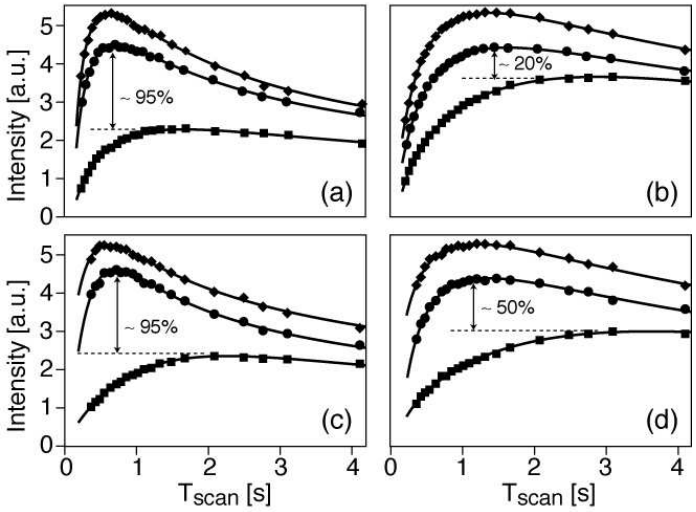


Figure 6.2 Average S/N ratios per unit time (intensity) plotted as a function of the scan time (pulse sequence duration plus recycle delay) obtained with the BEST-Jcomp HMQC sequence without z-filter (diamonds), with additional z-filter (circles), and using the SS-HMQC2 sequence of Wu and Bax (squares) for fully protonated ubiquitin at (a) 600 MHz and (c) 800 MHz 1H frequency, and for 75% partially deuterated SiR-FP18 at (b) 600 MHz and (d) 800 MHz 1H frequency. Intensities were obtained by adding the peak intensities of all detected amide (excluding NH2) correlation peaks observed in the 2D spectra.

formance. A time-reversed EBURP2 pulse (EBURP2*), initially optimized for flip-back purposes, is used for ^{1}H excitation. As illustrated by the Bloch simulations in Figure 6.1b, EBURP2* leaves chemical shift evolution (and spin-spin coupling) active during half of the pulse duration thereby creating a linear phase gradient. This phase gradient is refocused during the subsequent 90° pulse applied with an EBURP2 shape and a 90° phase shift with respect to the EBURP2* pulse. An in-phase H_y coherence is created for the component that has not evolved under dipolar coupling. At the same time the EBURP2 creates inphase H_x coherence from H_z, and flips back orthogonal transverse coherence along the z-axis (see Figure 6.1b), thereby converting the antiphase coherence $2H_x^AH_z^B$ to $-2H_z^AH_x^B$. The combination EBURP2*-Δ-180°-Δ-EBURP2 allows efficient use of the relatively long pulse durations (on the order of 2 ms), and avoids general rotation 90° pulses that are known to be more sensitive to B_1-field inhomogeneities and pulse imperfections.

Figure 6.2 illustrates the effect of the BEST modifications on the sensitivity of the experiment for two protein samples at ^{1}H frequencies of 600 and 800 MHz: (a, c) 76-residue fully protonated ubiquitin and (b, d) a 167-residue fragment of the *Escherichia coli* sulfate reductase (SiR-FP18) deuterated at a level of 75% (Sibille et al. (2005)). 2D BEST-HMQC2 spectra were recorded without the J-mismatch

compensation filters setting the effective evolution delays Δ_1 and Δ_2 to 26 ms. The average sensitivity (intensity per unit time) of BEST-HMQC2 is plotted as a function of the scan time (sum of pulse sequence duration and recovery delay), and compared to results obtained using the SS-HMQC2 sequence (Wu and Bax (2002), figure 10.1) recorded under comparable conditions. As a result of the selective ^{1}H manipulation in BEST-HMQC2 the maximum in the sensitivity curves shifts towards shorter scan times, reflecting the significant decrease in effective longitudinal relaxation times. For protonated ubiquitin, scan times of ~600 ms are optimal, whereas for the highly deuterated sample of SiR-FP18 a scan time of ~1.5 s provides highest S/N ratios. These optimal repetition rates depend only little on the magnetic field strength. An average sensitivity gain of a factor of ~2, and individual gains of up to a factor of 4 (data not shown) are observed for ubiquitin if the scan times of both experiments are optimized independently. Even for the deuterated sample of SiR-FP18, an average sensitivity gain of 20% at 600 MHz and 50% at 800 MHz is obtained. This observation can be explained by the fact that most methyl groups remain protonated at least at one of the three methyl hydrogen sites. ^{1}H-^{1}H spin diffusion to methyl groups then still allows quite efficient energy dissipation at these sites of high mobility (methyl rotation). Additionally, and in contrast to the SS-HMQC2 experiment, water polarization is left unperturbed rather than dephased, which also contributes to faster effective relaxation of the exchangeable amide protons. Even higher sensitivity gains are observed in the fast pulsing regime when using short recycle delays for both SS-HMQC2 and BEST-HMQC2 experiments. This may be relevant for situations where short overall acquisition times become important, e.g. for sample stability reasons.

6.3.2 J-mismatch compensation

Each of the two HMQC building blocks in the sequence of Figure 6.1a requires the adjustment of two transfer delays $\tau = 1/(2J_{NH})$, where J_{NH} is the heteronuclear coupling constant between amide ^{1}H and ^{15}N. While the scalar J couplings are quite uniform along the polypeptide chain with very little variation from one amide site to another, in weakly aligned molecules the RDC adds to the J coupling resulting in effective spin coupling constants $J^{eff}_{NH} = J_{NH} + D_{NH}$ that cover a range of typically ± 20-40% with respect to the mean value. It is therefore no longer possible to adjust the τ delays for all amide sites simultaneously, resulting in a sensitivity loss that depends on the degree of J-mismatch $f^{mis}_{NH} = 2\tau J^{mis}_{NH}$.

In a standard HMQC experiment the main coherence transfer pathway (P_I) is as follows:

$$H_y \xrightarrow{J^{eff}_{NH}} 2H_xN_y(t_1) \xrightarrow{J^{eff}_{NH}} H_y \sin^2(\pi f^{mis}_{NH}/2) \tag{6.3}$$

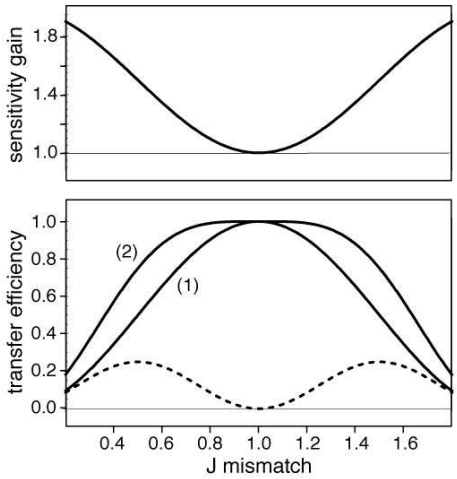


Figure 6.3 Simulations of ^{1}H-^{15}N transfer efficiencies (see Eqs. 6.3 and 6.4) in HMQC-type experiments with (2) and without (1) the additional J-mismatch compensation elements of Figure 6.1(a). The transfer efficiency of the additional coherence transfer pathway (P_{II}) created by the J-mismatch compensation elements is plotted as a dashed line. In addition, the expected signal gain for a single HMQC block as a function of J-mismatch is shown on top.

The efficiency of the heteronuclear transfer steps, given by $\sin^2(\pi f_{NH}^{mis}/2)$ neglecting spin relaxation, translates into a signal loss of $\approx$ 20% for a J-mismatch of f_{NH}^{mis} = 0.7 or 1.3 (see Figure 6.3). In order to increase the performance of the HMQC sequence for partially aligned protein samples, J-mismatch compensation elements (denoted (E1) and (E2) in the sequence of Figure 6.1a) are inserted. The J-mismatch compensation technique used here is conceptually similar to previously proposed broadband INEPT schemes (Nielsen et al. (1989); Wimperis and Bodenhausen (1986)). J-mismatch compensation allows recovering part of the "lost" magnetization by creating a second coherence transfer pathway (P_{II}) that adds to the detected NMR signal:

$$H_y \xrightarrow{J_{NH}^{eff}} 2H_xN_x(t_1) \xrightarrow{J_{NH}^{eff}} H_y \cos^2(\pi f_{NH}^{mis}/2)\sin^2(\pi f_{NH}^{mis}/2) \tag{6.4}$$

The transfer efficiency in J-mismatch compensated HMQC is then given by the sum of the two pathways P_I and P_{II}. In the J-mismatch range $0.7 < f_{NH}^{mis} < 1.3$ the ^{1}H-^{15}N transfer becomes almost independent of the effective coupling constant J_{NH}^{eff} with transfer efficiencies of more than 0.95 over the whole range (Figure 6.3). The amide ^{1}H-^{1}H coupling evolution is not affected by these additional pulse sequence elements because no 90° ^{1}H pulse is required. The J-mismatch compensation elements can thus be inserted in the ^{1}H-^{1}H coupling evolution periods of the BEST-HMQC blocks. Therefore, no additional delays are required and the im-

proved heteronuclear transfer efficiency of the HMQC blocks directly translates into intensity gains for the cross- and diagonal peaks depending on the J-mismatch of the involved amide groups. The only price to pay for the J-mismatch compensation, besides a few additional pulses, is a reduction (by ∼11 ms) in the available maximal t_1 and t_2 evolution times. This however does not present a strong limitation as long as transfer delays Δ_1, $\Delta_2 > 30$ ms are chosen for the amide ^{1}H-^{1}H coupling evolution.

6.3.3 Artefact suppression using an additional z-filter

To suppress undesired antiphase $2H_zH_x$ coherences before detection, while leaving inphase H_y coherence unperturbed, an optional z-filter element can be used. In short, the filter works as follows: the inphase H_y coherence is stored along the z-axis by a 90° EBURP2* pulse while antiphase $2H_xH_z$ coherence is converted to zero-quantum (ZQ) and double-quantum (DQ) coherences by the same pulse. The DQ coherence is destroyed by a pulsed field gradient. The ZQ coherence, that remains unaffected by the gradient pulse, evolves during a delay $1/(2J_{NH})$ into $8H_+H_-\ N_zN_z$-type coherence that is not detected in the presence of ^{15}N decoupling during data acquisition. The use of a high-performance ^{15}N composite decoupling sequence, such as WALTZ-16, is important to avoid partial detection of undesired coherence transfer pathways resulting in phase distortions in the ^{1}H dimension. The subsequent EBURP2 pulse then converts H_z to H_y for final detection. Finally, a WATERGATE sequence (Piotto et al. (1992)) is added for water suppression purposes.

6.4 Experimental results

The performance of the BEST-Jcomp-HMQC2 experiment has been tested on a 1.25 mM sample of fully protonated 15N-labeled ubiquitin. To create a weak alignment in the magnetic field, the ubiquitin sample was dissolved in a 15% w/v DMPC/DHPC mixture at a molar ratio of 3:1. The pH of the sample was adjusted to 6.6 in a 10 mM phosphate buffer. The resulting molecular alignment tensor, characterized on the basis of a set of D_{NH} coupling constants measured from IPAP-HSQC spectra (Andersson et al. (1998); Ottiger et al. (1998)) and the ubiquitin solution structure (pdb entry 1D3Z), showed an axial component of A_a = 17.8· 10^{-4} and a rhombic component of A_r = 2.6· 10^{-4}. Under these alignment conditions the J-mismatch for individual amide sites ranged from 0.6 to 1.3 assuming a scalar coupling constant J_{NH} = 92 Hz. All spectra were acquired at 800

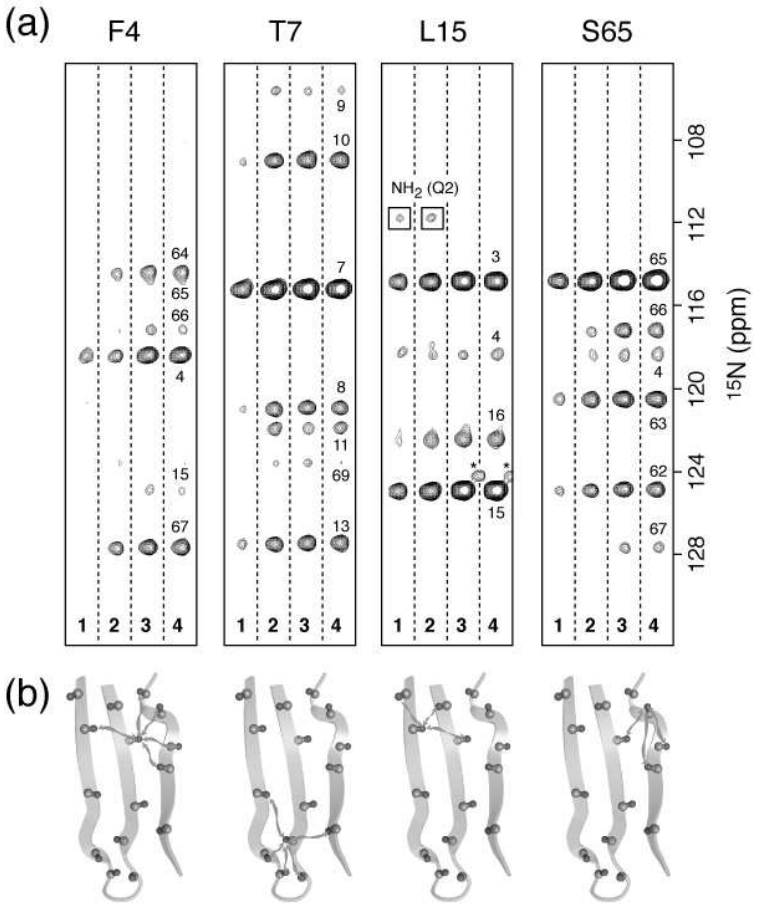


Figure 6.4 2D ^{15}N (ω_1)-^{15}N (ω_2) strips extracted at the amide ^{1}H (ω_3) frequency of several residues in the β-sheet of ubiquitin. 3D data sets were recorded at 800 MHz (28° C) on a sample of ^{15}N-labelled, fully protonated ubiquitin dissolved in a DMPC/DHPC mixture. The results obtained with the following pulse sequences are shown: (1) SS-HMQC2, (2) BEST- HMQC2 without J-mismatch compensation and z-filter, (3) BEST-Jcomp-HMQC2 with additional J-mismatch compensation, and (4) BEST-Jcomp-HMQC2 with J-mismatch compensation and z-filter. For all experiments the ^{1}H-^{1}H transfer delays (corresponding to Δ_1 and Δ_2 in the sequence of figure 6.1a) were adjusted to 41 ms. 43 (t_1) $\times$ 43 (t_2) $\times$ 400 (t_3) complex points were recorded for spectral widths of 2400 Hz (ω_1), 2400 Hz (ω_2) and 10000 Hz (ω_3). Optimal recycle delays were used for all experiments: 450 ms for BEST-Jcomp-HMQC2, and 1.4 s for SS-HMQC2. The total acquisition time was 10 hours for the BEST experiments, and 24 hours for the SS-HMQC2 spectrum. The intensities in the SS-HMQC2 spectrum were scaled by a factor 0.65 to account for the unequal experimental times. A star indicates a residual peak from another ^{1}H plane (residue). (b) Schematic drawing of the β-sheet structure in ubiquitin. The arrows indicate amide pairs for which RDCs could be measured from the spectral regions displayed in (a), illustrating that important long-range structural information is obtained from these data.

MHz ^{1}H frequency and 28°C sample temperature on a Varian INOVA spectrometer equipped with a cryogenic triple-resonance probe and shielded z-gradients. 3D spectra were recorded using different options of the pulse sequence displayed in Figure 6.1a: with/without J-mismatch compensation, and with/without z-filter. In addition, a SS-HMQC2 spectrum (Wu and Bax, 2002) was acquired for comparison using identical acquisition parameters except for the interscan delay, which was set to 450 ms in BEST- HMQC2 and to 1.4 s in SS-HMQC2 for optimal sensitivity in both experiments. Examples of 2D (ω_1, ω_2) strips extracted from these data sets are shown in Figure 46.4a. On average, a sensitivity gain of approximately a factor of 2 is obtained from longitudinal relaxation enhancement when using BEST-Jcomp-HMQC2 instead of SS-HMQC2. This gain, however, is not uniform along the peptide chain and enhancement factors of 3 to 5 are observed for amide groups in the hydrophobic core of the protein characterized by a high local proton density and little internal dynamics. Interestingly, these are also the molecular regions where the most useful long-range structural information is obtained, as illustrated for the β-sheet structure of ubiquitin (Figure 6.4b). J-mismatch compensation yields significant further signal enhancements for amide groups with large D_{NH} couplings. Sensitivity gains of up to a factor of 2 are observed when adding the J-mismatch compensation elements for cross and diagonal peaks were both amide groups involved in the correlation show large J-mismatch. The combination of both effects yields enhancement factors of the S/N ratio for individual correlation peaks in the range of ~1.5 to 8. The increased sensitivity allowed us to quantify additional peaks and thus obtain more structural restraints as compared to the SS-HMQC2 experiment.

In addition to the 113 cross peaks that are observed in both experiments 41 additional cross peaks with intensities above a cutoff level of 3 times the noise level could be quantified. Exceptions are correlations between amide ^{1}H and the NH_2 groups of Asn and Gln side chains. These cross peaks have significantly reduced intensity in BEST-Jcomp-HMQC2 because of the additional J_{NH} coupling evolution of $2N_yH_x$-type coherence during the J-compensation pulse sequence elements (see residue L15 in Figure 6.4a). Figure 6.5a shows the correlation of D_{HH} values extracted from 3D spectra measured using the SS-HMQC2 and BEST-Jcomp-HMQC2 sequences, illustrating that the same level of accuracy is obtained from the two experiments, although with a significantly higher precision for the sensitivity-optimized BEST-Jcomp-HMQC2 version.

We have also experimentally evaluated the effect of the additional z-filter element on the overall sensitivity of the BEST-Jcomp-HMQC2 experiment and the accuracy of the measured amide ^{1}H-^{1}H RDCs. Overall the intensity of most correlation peaks recorded with the additional z-filter is slightly reduced, by about 10-15% in the case of the weakly aligned ubiquitin sample used for this study (Fig-

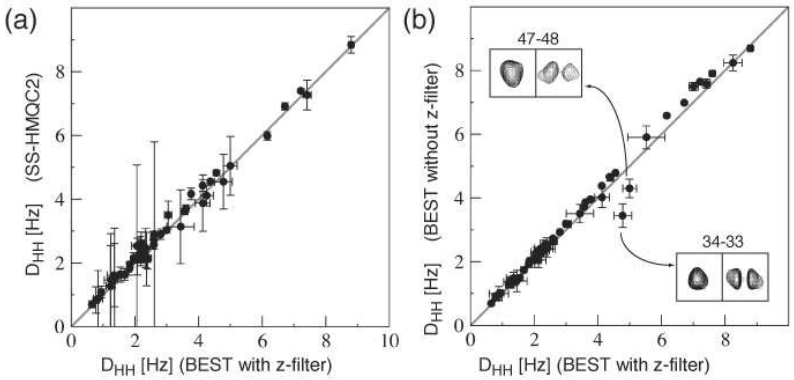


Figure 6.5 (a) Correlation plot of amide ^{1}H-^{1}H RDCs measured using the BEST-Jcomp-HMQC2 sequence of Figure 6.1a (x-axis) and SS-HMQC2 (Wu and Bax (2002)) (y-axis). The total acquisition time was 10 hours for BEST-Jcomp-HMQC2 and 24 hours for SS-HMQC2. (b) Correlation plot of amide 1H-1H RDCs extracted from 3D BEST-Jcomp-HMQC2 spectra recorded with (x-axis) and without (y-axis) the additional z-filter. The experimental error bars were estimated on the basis of intensity uncertainties of 1.5 times the noise level. (b) The absolute values of the D_{HH} coupling constants were extracted from a single spectrum. Therefore all D_{HH} couplings were assumed to be in the 0 to $1/2\Delta$=12.2 Hz range ($\Delta = \Delta_1 = \Delta_2$). The periodicity of the trigonometric functions in Eq. 6.1 does not allow the distinction between coupling constants D and $1/\Delta - D$. In order to solve this ambiguity a second data set recorded using a different transfer delay Δ is required. The inserts in (b) show ^{1}H line shape distortions observed for some cross peaks (right spectrum) that are completely removed by addition of the z-filter (left spectrum).

ure 6.4a), which is in agreement with the data shown in Figure 6.2. There are two possible sources of signal losses: first, and probably most importantly, ^{1}H longitudinal relaxation during the z-filter element and ^{1}H transverse relaxation during the WATERGATE sequence of ~2 ms duration; second, as can be seen from Figure 6.2, the optimal scan time is slightly shifted towards longer delays as compared to the experiment without the z-filter element, indicating that the additional pulses decrease the steady-state polarization of the aliphatic protons, and thus the sensitivity of the experiment. The small intensity increase observed for some cross peaks is explained by the removal of artifacts with negative intensity in the ω_3(^{1}H) dimension. The difference in the RDC values extracted from the two experiments recorded with and without z-filter is surprisingly small for most of the observed correlations (Figure 6.5b), indicating that also without the z-filter meaningful structural restraints can be obtained. Only a few peaks show significant distortions along the ^{1}H dimension (see inserts in Figures 6.5b) that result in erroneous peak intensities and thus stronger deviations in the measured DHH coupling constants. These peak distortions are absent in the spectra recorded with the additional z-filter illustrating that the antiphase $2H_xH_z$ coherence is efficiently suppressed. The

z-filtered version of the experiment thus yields higher accuracy at the expense of a small loss in sensitivity. However, the outliers are easily identified from inspection of the peak shape and accurate coupling constants can also be obtained from the experiment without the z-filter element if these peaks are manually removed from the analysis.

6.4.1 Conclusions

We have presented BEST-Jcomp-HMQC2, a new pulse sequence that allows accurate measurement of amide ^{1}H-^{1}H residual dipolar couplings in weakly aligned proteins. ^{1}H-^{1}H RDC measurements provide important long-range restraints for NMR structure determination. This information is available immediately after backbone resonance assignment without the need for any additional tedious side chain resonance assignment or NOE analysis. ^{1}H-^{1}H RDCs are therefore especially useful in the context of fast de novo fold elucidation using only a limited set of NMR-derived structural restraints. It has been shown recently (Bouvignies et al. (2006)) that high structural precision is obtained by combining local orientational information from RDCs between covalently-bound nuclei in the protein backbone with long-range translational and orientational information from amide ^{1}H-^{1}H RDCs. The high sensitivity provided by the BEST-Jcomp-HMQC2 experiment makes it attractive to measure amide ^{1}H-^{1}H RDCs on a routine basis for partially aligned protein samples without the need for (per)deuteration.

7 Combination of fast-pulsing techniques with alternative data sampling

The experiments introduced in previous chapters used standard sampling of indirect time domains, and achieved accelerated data acquisition by reducing the recovery delay between scans. Because of the need to sufficiently sample these time domains, the minimal experimental time of these techniques is limited by the number of scans.

Interestingly, fast-pulsing methods are compatible with alternative sampling schemes, and the next sections show the combination of the SOFAST HMQC approach with Hadamard- and spatial encoding. Both approaches allow the acquisition of site-resolved 2D information in experimental times of ca. 1 second.

7.1 Hadamard frequency-encoded SOFAST-HMQC for ultrafast two dimensional protein NMR

7.1.1 Introduction

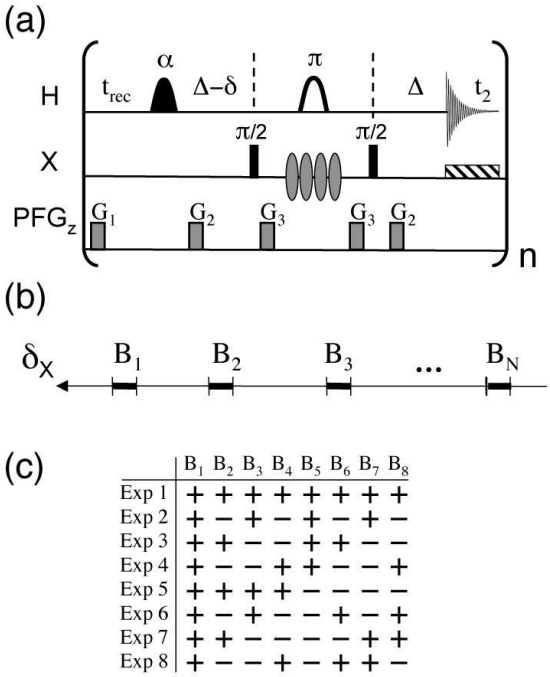
Multidimensional nuclear magnetic resonance (NMR) has proven a useful tool for the structural and dynamical characterization of biological macromolecules. Multidimensional NMR is also very powerful for real-time site-resolved studies of kinetic processes in these biomolecules. An important limitation for such applications, however, remains the intrinsically low time resolution provided by multidimensional NMR because each 2D spectrum requires at least a few minutes, each 3D spectrum a few hours of data acquisition to obtain an acceptable spectral resolution. Recently there has been an increasing interest in developing new methods for faster multidimensional data acquisition (for a review see Freeman and Kupce (2003)). The experimental time required to record a complete spectrum is determined by the number of repetitions of the basic pulse scheme (scan) and the duration of a single scan. The number of repetitions has to take into account phase cycles and quadrature detection, and it determines the spectral resolution in the individual indirect frequency dimensions. The length of a single scan is generally dominated by the time required for the spin system to relax towards its thermal equilibrium after a multiple pulse sequence (recycle delay). Most approaches proposed for fast NMR data acquisition aim at reducing the number of recorded data points. Examples are projection NMR (Brutscher et al. (1994); Kim and Szyperski (2003); Kupce and Freeman (2004a); Szyperski et al. (1993b)), Hadamard NMR (Brutscher (2004); Kupce et al. (2003)), single scan NMR (Frydman et al. (2002)), and non-linear data recording and processing (Hoch and Stern (2001); Mandelshtam (2000); Rovnyak et al. (2004a)). Alternatively, one may reduce the length of a single scan, generally dominated by the time required for the spin system to relax towards its thermal equilibrium after a multiple-pulse sequence (recycle delay). The sensitivity of certain pulse sequences for short recycle delays can be enhanced by the use of spin-lattice relaxation enhancement techniques (Atreya and Szyperski (2004); Pervushin et al. (2002)) or Ernst-angle excitation (Ernst et al. (1987); Ross et al. (1997)). Recently we have introduced SOFAST HMQC (Schanda and Brutscher (2005)), an experiment that combines the advantages of spin-lattice relaxation enhancement and Ernst-angle excitation to reduce inter-scan delays. SOFAST HMQC has proven to provide the required sensitivity to record well-resolved ^{1}H-^{15}N and ^{1}H-^{13}C correlation spectra of proteins at millimolar concentration within a few seconds of acquisition time. IPAP SOFAST HMQC, a slightly modified version of the originally proposed experiment, allows application on spectrometers

equipped with cryogenic probes (Schanda et al. (2005)). In the IPAP SOFAST HMQC experiment the heteronuclear decoupling during acquisition is replaced by an IPAP filter (Andersson et al. (1998); Ottiger et al. (1998)) thus reducing probe heating and duty cycle problems for the short inter-scan delays required in SOFAST HMQC. Here, we show that SOFAST HMQC combined with Hadamard-type frequency editing in the ^{15}N dimension allows to further reduce the minimal experimental time. As demonstrated for the protein ubiquitin this "hybrid" technique, Hadamard-encoded SOFAST HMQC or IPAP SOFAST HMQC, provides site-specific resolution for a large number of nuclei in the protein within one second of data acquisition. This is, to the best of our knowledge, the first time that such 2D spectral information is obtained on a fairly concentrated protein sample (2 mM) in about one second of experimental time using a standard high-field NMR spectrometer. The increased time resolution provided by the Hadamard-encoded SOFAST HMQC opens the way to real-time investigations of protein kinetics with characteristic time constants of less than a second.

7.1.2 Hadamard-encoded SOFAST HMQC

The pulse sequence for Hadamard-encoded SOFAST-HMQC is shown in figure 7.1a. The basic features of Hadamard-encoded SOFAST-HMQC are the same as discussed previously for SOFAST-HMQC (Schanda and Brutscher (2005)) except for the ^{15}N time-domain frequency labeling which is replaced by a Hadamard encoding scheme. Therefore only the main features will be summarized shortly. First, ^{1}H pulses are applied band-selectively which ensures that the unobserved protons remain in their equilibrium state and provide a thermal bath of proton spin polarization which enables faster spin-lattice relaxation of the excited ^{1}H spins via dipolar interactions. Second, the first ^{1}H pulse has an adjustable flip angle α that allows further optimization of the sensitivity of the experiment for a chosen (short) scan time. In practice, a flip angle $90^\circ \leq \alpha \leq 160^\circ$ is chosen to ensure that part of the ^{1}H magnetization is restored along the z-axis by the following 180° pulse. Third, the small number of rf pulses reduces signal loss due to pulse imperfections and B_1 field inhomogeneities, and limits the effects of sample and probe heating. The application of only 2 (band-selective) ^{1}H pulses also ensures minimal perturbation of the unobserved proton spin polarization, thus optimally exploiting the spin-lattice relaxation enhancement effect discussed above.

In the sequence of Fig. 7.1a the free evolution delay t_1 of the standard SOFAST-HMQC is replaced by a Hadamard-type ^{15}N frequency encoding. In Hadamard spectroscopy (Kupce et al. (2003)) N different frequency "channels" or "bands" are defined (Fig. 7.1b) and each band is manipulated individually by means of frequency-selective radio-frequency (rf) pulses.



	B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8
Exp 1	+	+	+	+	+	+	+	+
Exp 2	+	−	+	−	+	−	+	−
Exp 3	+	+	−	−	+	+	−	−
Exp 4	+	−	−	+	+	−	−	+
Exp 5	+	+	+	+	−	−	−	−
Exp 6	+	−	+	−	−	+	−	+
Exp 7	+	+	−	−	−	−	+	+
Exp 8	+	−	−	+	−	+	+	−

Figure 7.1 (a) Hadamard-encoded SOFAST-HMQC pulse sequence to record fast ^{1}H-^{15}N correlation spectra of proteins. The flip angles of the individual rf pulses, all applied along the xaxis, are given above the pulse symbols. Shaped pulse symbols indicate the use of bandselective pulses. The pulse with variable flip angle α has a polychromatic PC9 shape (Kupce and Freeman (1994)) and band-selective ^{1}H refocusing is realized using a REBURP profile (Geen and Freeman (1991)). The transfer delay Δ is set to $1/2J_{HN}$, the delay δ accounts for spin evolution during the PC9 pulse, and is adjusted to yield pure phase spectra in the ^{1}H dimension. t_{rec} is the recycle delay between scans. Adiabatic WURST-2 decoupling (Kupce and Freeman (1996)) is applied on X during detection. Hadamard ^{15}N frequency labeling is realized using band-selective refocusing pulses with a sinc profile. The phase of these pulses is interchanged between x and y to realize (+) and (-) encoding, respectively, according to the appropriate Hadamard matrix. The Hadamard encoding pulses were generated by vector addition of the individual band-selective pulse shapes (Kupce and Freeman (1993)). (b) For Hadamard encoding N different (non-overlapping) frequency bands are defined in the ^{15}N spectrum. N experiments need to be recorded with a sign encoding according to a Hadamard matrix of order N. The same matrix is then used to disentangle the NMR signals from the individual frequency bands during data processing. The Hadamard matrix for the case N=8 is shown in (c). The all-(+) frequency band (B_0) is prone to experimental artifacts and is not used. Therefore only N-1 frequency bands have to be defined, although N experiments are recorded. If necessary the experiment may be repeated n-times in order to increase the signal to noise ratio.

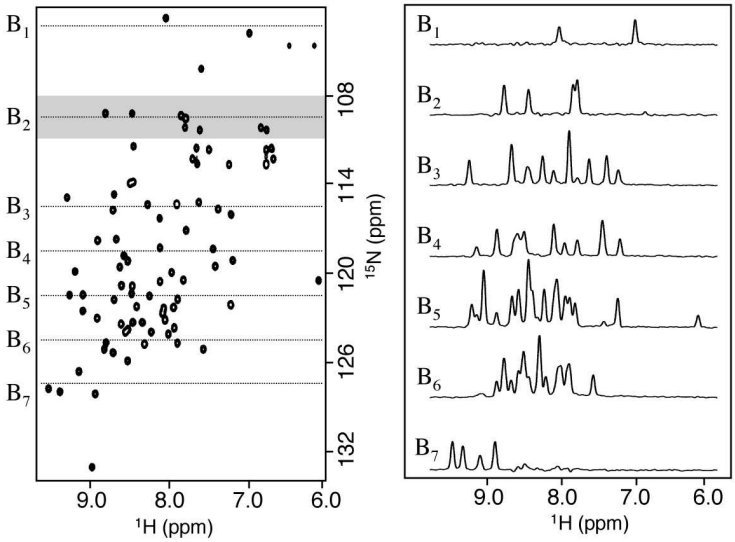


Figure 7.2 (a) Definition of the ^{15}N frequency bands used for the Hadamard-encoded SOFASTH-MQC spectrum of ubiquitin shown in (b). The center positions of the individual ^{15}N frequency bands are indicated by dotted lines. The band width of about 3 ppm covered by the Hadamard encoding pulses used for recording spectrum (b) is exemplified for band B_2. The spectrum in (a) has been recorded with a standard ^{1}H-^{15}N HSQC-sequence as provided by the Varian BioPack. The spectrum in (b) has been recorded on a 600MHz spectrometer on a 2mM sample of ubiquitin (pH 6.2, 25°C) using the pulse sequence of figure 7.1a with the following parameter settings: α=130°, Δ =5.4 ms, δ = 1.6 ms, t_{rec} = 10 ms, t_2^{max} = 40 ms, N = 8, n = 64 (+ 4 dummy scans). The band-selective ^{1}H excitation (PC9) and refocusing pulses (REBURP) were centered at 8.0 ppm covering a bandwidth of 4.0 ppm, resulting in pulse lengths of 3.0 ms (PC9) and 2.03 ms (REBURP). The Hadamard encoding pulses were applied with a sinc shape truncated after the first side lobe and a pulse length of 22ms covering an effective band width of about 180 Hz (3 ppm). ^{15}N decoupling during t2 was realized using WURST-2 Kupce and Freeman (1996) at an average field strength of $\gamma B_1/2\pi$= 550 Hz. The experimental time was 1 minute.

One may distinguish between frequency-selective Hadamard NMR spectroscopy, where the band width covered by the selective encoding pulses equals the line width of the individual resonances in the Hadamard encoded frequency dimension, or band-selective Hadamard spectroscopy, where larger band widths are chosen resulting in the projection of all resonance lines within this band on a single point along the Hadamard dimension. The sign of the NMR signal originating from the N bands is changed according to a Hadamard matrix of order N (Fig. 7.1c). This sign encoding ensures that each frequency band contributes fully to the detected signal in each of the experimental repetitions, thus providing the same multiplex advantage as standard time-domain NMR spectroscopy. N experiments need to be recorded to separate the individual bands during the processing stage

by an inverse Hadamard transformation. The main interest of Hadamard spectroscopy is that the number of bands, and thus the number of repetitions of the experiment, can be freely adjusted without compromising spectral resolution, as long as a Hadamard matrix exists for this particular choice (Hadamard matrices exist for orders N=4j and N=2j with j an integer).

Hadamard frequency encoding in the sequence of figure 7.1a is realized by band-selective ^{15}N refocusing pulses with phase +x for (+) encoding and +y for (-) encoding. The individual band-selective refocusing pulses are combined to a single shape by vector addition (Kupce and Freeman (1993)), and a different shaped pulse is applied for each of the N experiments. In addition, a ^{1}H 180° pulse is applied simultaneously to the Hadamard encoding pulses to refocus ^{1}H chemical shift evolution. The heteronuclear J_{NH} coupling evolution, however, is not refocused by this 180° pulse sandwich. For NH spin systems the multiple quantum coherence remains unaffected by scalar J_{NH} coupling evolution during the Hadamard encoding. For NH_2 spin systems, however, the peak intensities in the Hadamard-encoded SOFAST-HMQC spectrum are modulated as a function of the encoding time. Proper adjustment of the delay between the two 90° ^{15}N pulses thus allows filtering out signals from NH_2 spin systems. An example of a Hadamard-encoded SOFAST-HMQC experiment is shown in figure 7.2 for the small protein ubiquitin (76 residues). From a standard ^{1}H-^{15}N correlation map seven ^{15}N frequency bands were selected, distributed over the whole frequency range as shown in Fig. 7.2a. An additional "dummy" band was assigned to the all-(+) frequency band (B_0) that contains signals from ^{15}N spins whose chemical shifts are not or only partially refocused by the Hadamard encoding pulses (see figure 7.3).

7.1.2.1 Choice of Hadamard encoding pulses

Our principal aim was to detect as many peaks as possible in the Hadamard-encoded ^{1}H-^{15}N correlation spectrum without too much of peak overlap in the ^{1}H dimension. Therefore, 5 bands were equally spaced along the ^{15}N dimension ($\Delta\nu$ =190 Hz) in the most crowded spectral region, and 2 additional bands were defined in the down-field region comprising only few resonances. Different band-selective refocusing pulse shapes were tested to realize band-selective Hadamard encoding of about 3 ppm band width (corresponding to 180 Hz on a 600 MHz spectrometer). Bloch simulations of the spin evolution during the Hadamard ^{15}N encoding using Gaussian, sinc or rSNOB shapes are shown in figure 7.3. The best compromise between short pulse length and high selectivity was obtained from sinc pulses, truncated after the first side lobe.

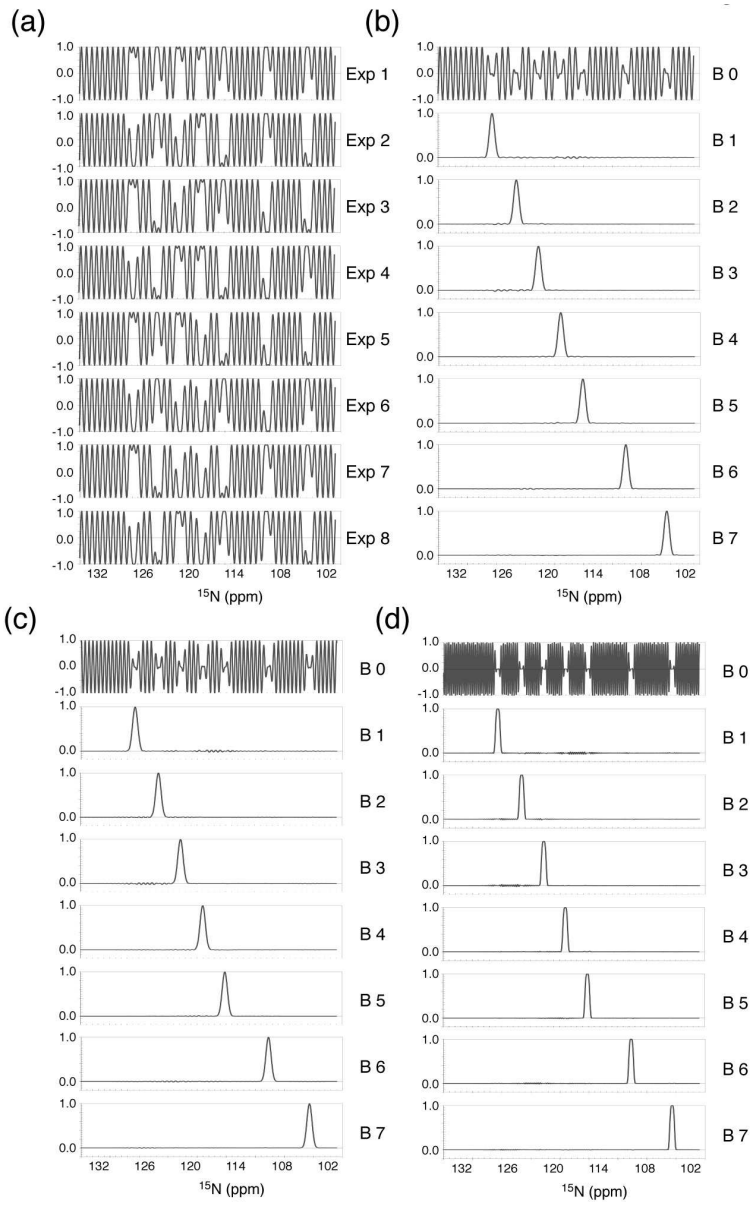


Figure 7.3 Bloch simulations of the spin evolution during Hadamard ^{15}N encoding based on the frequency bands defined in figure 7.2. The calculations, performed using the Varian "Pulsetool", assume an initial N_y state of the spin system and neglect spin relaxation effects. The N_y component present after the Hadamard encoding pulses is plotted. (a) Refocusing performance of the Hadamard pulses used for the different experiments (Exp 1 - Exp 8) as a function of the ^{15}N frequency for a sinc shape (1 side lobe) of 22 ms length. (b) Same simulations as in (a), but the plots correspond to the results after decoding with the appropriate Hadamard matrix to separate the signals from the individual bands (B0 - B7). (c) Results obtained for a Gaussian pulse shape (truncation level 0.01) of 30 ms length. (d) Results obtained for a r-SNOB pulse shape (Kupce et al. (1995)) of 62 ms length.

7.1.3 Application to ubiquitin

The result of a Hadamard-encoded ^{1}H-^{15}N SOFAST-HMQC experiment performed with sinc-encoding pulses (pulse length of 22 ms at 60MHz ^{15}N frequency) and N=8 is shown in Fig. 7.2b. The spectrum was recorded on a 2mM sample of ubiquitin (25°C, pH 6.2) on a 600MHz Varian INOVA spectrometer equipped with a standard (non cryogenic) triple resonance probe. The series of 1D spectra corresponds to projections of the 2D spectrum (Fig. 7.2a) along the ^{15}N dimension over the individual bands. No signals from the NH_2 groups in Asn and Gln side chains are detected in this spectrum, because they are filtered out by the J_{NH} coupling evolution during the chosen Hadamard encoding pulse length. Repeating the whole experiment 64 times (n) increased the signal to noise ratio, and allowed to evaluate the spectral quality and to detect artifacts induced by this particular Hadamard encoding. As can be appreciated from Fig. 7.2b the individual spectra show little artifacts, and correspond to quite nicely resolved 1D amide proton spectra of ubiquitin.

Once the feasibility of our method demonstrated, we have repeated the same experiment in only about 1 second of acquisition time by setting n = 1. The resulting spectrum is shown in figure 7.4. Despite the reduced signal to noise ratio, because of the short overall experimental time, most of the amide protons in ubiquitin are visible in this spectrum, and a significant number of them is spectrally resolved. Overall we count from the spectrum in figure 7.4 a total of 42 resolved peaks, 11 peaks corresponding to 2 overlapping resonances, and an additional peak where 3 resonances superpose. This clearly demonstrates the potential of Hadamard-encoded SOFAST-HMQC for obtaining residue-specific information on a large number of amide sites within one second of experimental time.

7.1.4 Conclusions

In summary, we have shown that site-resolved spectral information can be obtained within one second of acquisition time on a ^{15}N labeled sample of a small protein at millimolar concentration using a high field NMR spectrometer and Hadamard encoded SOFAST-HMQC. The experiment can be designed to focus on only a few sites of particular interest using frequency-selective Hadamard encoding, or cover most of the spectral range, as demonstrated here, using band-selective Hadamard encoding. The use of higher field magnets will allow increasing the selectivity of the Hadamard encoding without increasing the lengths of the encoding pulses. It is also possible to adjust the width for each band individually in order to optimize the number of resolved amide peaks. Higher selectivity will be im-

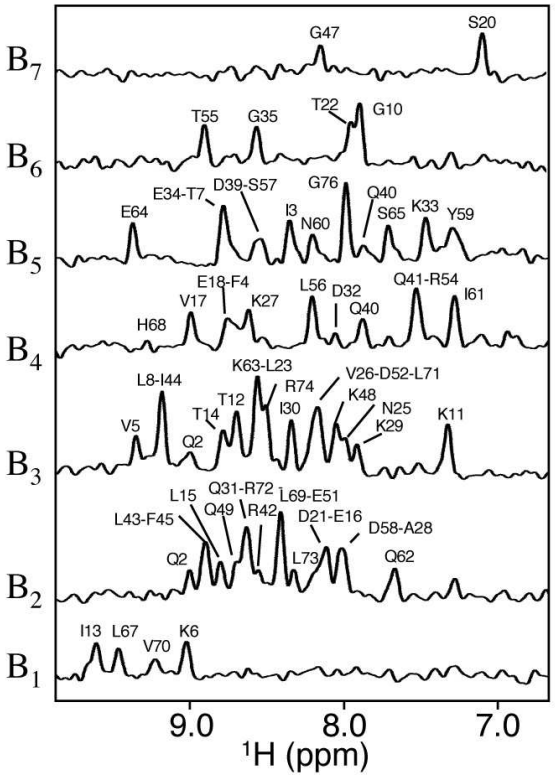


Figure 7.4 Ultrafast Hadamard-encoded SOFAST-HMQC spectrum of ubiquitin recorded with the pulse sequence of figure 7.1(a). The same frequency bands and acquisition parameters were used as described in figure 7.2, except for the number of repetitions which was set to N = 8, n=1 (+ 4 dummy scans) resulting in an experimental time of about 1 second. Assignments (residue number and type) are shown for the detected amide correlation peaks. 67 amide resonances are detected in this spectrum, out of the 70 cross peaks observed in the standard HSQC spectrum shown in fig. 7.2a.

portant for application to larger proteins in order to resolve individual amide sites in the Hadamard-encoded spectra. The reduced sensitivity, because of relaxation losses during the longer Hadamard encoding pulses, may be compensated by the use of cryogenic probes and Hadamard encoded IPAP SOFAST HMQC. The experiment is equally well suited for fast recording of ^{1}H-^{13}C correlation spectra of proteins. Hadamard-encoded SOFAST HMQC provides a new tool for following the changes in peak intensities occurring during some kinetic event with a seconds resolution for a selected number of nuclear sites. Interesting examples are the measurement of H/D exchange rates, or the measurement of some slow binding kinetics. It is less suited for monitoring processes that do not only change the peak intensities but also their positions, as frequency changes in the ^{15}N dimension

within one band do not translate into changes in the Hadamard-encoded spectra. Hadamard-encoded SOFAST-HMQC is only one example of what one could call ”hybrid” fast acquisition methods. It is equally possible to combine SOFASTH-MQC with spatial frequency encoding (see section 7.2) or the use of non-linear data acquisition and processing schemes. If the intrinsic sensitivity of NMR spectrometers continues to increase as observed over the last decade, one may expect that such (ultra-) fast acquisition schemes will soon become applicable to a wide range of biomolecular systems.

7.2 UltraSOFAST HMQC NMR and the Repetitive Acquisition of 2D Protein Spectra at a Rate of a second

7.2.1 Abstract

Following unidirectional biophysical events such as the folding of proteins or the equilibration of binding interactions, requires experimental methods that yield information at both atomic-level resolution and at high repetition rates. Towards this end a number of different approaches enabling the rapid acquisition of 2D NMR spectra have been recently introduced, including spatially-encoded ultrafast 2D NMR spectroscopy and SOFAST HMQC NMR. Whereas the former accelerates acquisitions by reducing the number of scans that are necessary for completing arbitrary 2D NMR experiments, the latter operates by reducing the delay between consecutive scans while preserving sensitivity. Given the complementarities between these two approaches it seems natural to combine them into a single tool, enabling the acquisition of full 2D protein NMR spectra at high repetition rates. We demonstrate here this capability with the introduction of ultraSOFAST HMQC NMR, a spatially-encoded and relaxation-optimized approach that can provide 2D protein correlation spectra at $\sim$1 sec repetition rates for samples in the $\sim$2 mM concentration range. The principles, relative advantages and current limitations of this new approach are discussed, and its application is exemplified with a study of the fast hydrogendeuterium exchange characterizing amide sites in Ubiquitin.

7.2.2 Introduction

Multidimensional nuclear magnetic resonance (NMR) spectroscopy is a widely used tool in the study of biomolecular dynamics. Whereas fluctuations in proteins and nucleic acids on a pico- to millisecond timescale are best studied by relaxation measurements, residual dipolar couplings or spectral line shapes (Blackledge (2005); Mittermaier and Kay (2006); Palmer (2004)) unidirectional dynamic changes such as the exchange of amide protons for deuterons (H/D exchange) or the progression of an unfolded ensemble towards a folded structure, can be more naturally followed by real-time NMR spectral changes (Dempsey (2001); Dobson and Hore (1998); Zeeb and Balbach (2004)). An important limitation to such real-time measurements comes from the fact that conventional 1D NMR spectroscopy lacks the resolution needed to cope with multiple overlapping lines; on the other hand, the 2D NMR methods that are capable of resolving the individual atomic peaks arising from a macromolecule are intrinsically time-consuming. These longer ac-

quisition times come as a result of the manner by which the data, and in particular the indirect-domain spectral components, are sampled in 2D NMR (Aue et al. (1976); Jeener (1971)) Whereas the t_2-domain is directly monitored in 2D acquisitions by a single-transient physical digitization, the indirect-domain evolution is recorded by repeating the pulse sequence numerous times in association with increments of a corresponding delay parameter t_1. Given the identical Nyquist criteria defining bandwidth and resolution along both time-domains this means that, even if sensitivity were sufficient to monitor the data being sought in a single scan, multiple repetitions will still have to be performed. The total experimental time will in such cases be given by the product of the number of scans N_{scans} required for appropriately sampling the indirect domain, times a single-scan duration T_{scan} usually dominated by the need for allowing spins to relax back to thermal equilibrium prior to repeating an additional measurement. In contemporary protein NMR spectroscopy this recycle delay is associated with a ^{1}H spin-lattice relaxation time T_1 on the order of seconds; this in turn translates into acquisition times on the order of minutes even for the most sensitive kinds of 2D NMR experiments. Very rapid repetition rates on the order of Hz, which might be of great interest for the study of fast biomolecular events, are therefore unavailable to standard real-time 2D NMR acquisitions.

Driven partly by this need recent years have witnessed an increased interest in accelerating the acquisition of 2D NMR data, and to this end several methods have been proposed (Freeman and Kupce (2003)). In general terms these methods can be catalogued as acting either by reducing the T_{scan} repetition delay between the various increments, or by decreasing the overall number N_{scans} of scans needed to retrieve the spectra. The former group aims at maximizing sensitivity while allowing for very short recycle delays; this can be achieved by accelerating the spin-lattice relaxation of the spins of interest, (Diercks et al. (2005); Pervushin et al. (2002)) and in certain specific experiments by relying on optimized flip-angles (e.g., the Ernst angle) to enhance the steady-state magnetization of the excited spins (Ross et al. (1997)). Moreover, proposals have been recently made that combine these two features into single 2D and 3D NMR protocols (Schanda and Brutscher (2005); Schanda et al. (2005, 2006)), including the band-Selective Optimized Flip-Angle Short-Transient HMQC (SOFAST HMQC) NMR experiment. SOFAST HMQC enables a reduction of T_{scan} down to $\lesssim$100 ms while preserving high sensitivity, thus allowing for the recording of conventionally-sampled 2D ^{1}H-^{15}N or ^{1}H-^{13}C correlation spectra within minimal experimental times of just several seconds. Important among the group of experiments whose aim is to reduce the total number of necessary N_{scans} count accordion and projection NMR methods (Bersch et al. (2003); Ding and Gronenborn (2002); Kim and Szyperski (2003); Kupce and Free-

man (2004a); Szyperski et al. (1993b) Hadamard spectroscopy (Brutscher (2004); Kupce et al. (2003)), and other ingenious schemes that carry a nonlinear sampling of data points together with an appropriate data processing (Bruschweiler and Zhang (2004); Hoch and Stern (2001); Mandelshtam (2000); Rovnyak et al. (2004a)). Arguably, the most dramatic reduction in the number of required scans is achieved by ultrafast NMR; a protocol which allows one to record full, arbitrarily-high multi-dimensional NMR data sets within a single transient (Frydman et al. (2003, 2002); Shrot and Frydman (2003)). At the heart of this approach lies replacing the stepwise incrementation of the temporal parameter t_1 involved in 2D NMR, by an analogous encoding of the indirect-domain evolution along a spatial dimension z. This procedure can be carried out in a variety of different ways, usually involving the combined application of pulsed field gradients and frequency-swept spin manipulations, so as to replace the $\Omega_1 t_1$ time encoding by an analogous $C\Omega_1(z - z_0)$ spatial encoding with $C \propto t_1^{max}$ (sample length) a constant under our control and z_0 denoting an arbitrary spatial origin. The spatial winding of the spins entailed by this phase expression is preserved, just as in traditional 2D NMR experiments, by a coherent mixing scheme, to be subsequently unwound and read out over the course of the data acquisition using pulsed field gradients. This reading process can be repeated numerous times by oscillating the sign of the readout field gradients, allowing one to monitor a set of indirect-domain spectra as a function of the direct-domain t_2 evolution. Fourier transformation of the resulting data along t_2 thus provides the full $I(\Omega_1,\Omega_2)$ spectral information being sought, within a single transient.

Although the spatially-encoded NMR principles just described can yield 2D NMR spectra within a single scan, a full dynamic characterization will in general entail collecting a large number of transients. This demand will be intrinsic to real-time methods that look at unidirectional biomolecular processes by monitoring the full range of 2D NMR spectral changes occurring as a function of time. Moreover, even if ultrafast NMR makes the acquisition of 2D data feasible within a single scan, practical sensitivity and solvent suppression considerations often demand the signal averaging of at least a minimal number of transients. Indeed contemporary NMR hardware limits the per-scan sensitivity of ultrafast 2D acquisitions to the low-mM range (Shapira et al. (2004b)); these concentrations may exceed those typically encountered in biomolecular measurements attempting to detect transient chemical species, implying that several scans may have to be averaged anyhow to achieve an acceptable signal-to-noise ratio (SNR). Bearing these needs in mind it becomes clear that, even when considering the use of single-scan 2D NMR within a real-time biomolecular dynamic setting, this approach could also benefit from protocols that shorten the repetition delay T_{scan} while preserving maximum

SNR. The present study explores such possibilities by demonstrating that, if the very short inter-scan delays afforded by the SOFAST protocol are merged with the single-scan capabilities of ultrafast NMR, full 2D protein correlation spectra can be repeatedly retrieved at unprecedented rates. A number of provisions to be taken into account for maximizing the efficiency of the resulting ultraSOFAST NMR experiment are discussed, its ability to acquire real-time 2D NMR spectra of proteins undergoing a uni-directional dynamic process at mM concentrations within 1-2 sec experimental times is illustrated, and its benefits and drawbacks vis-à-vis other recent proposals in the field are compared.

7.2.3 Methods

Different variants could be considered for combining the single-scan capabilities of ultrafast NMR with the rapid repetition benefits of SOFAST protocols. The pulse sequence chosen to test the general performance of such spatially-encoded 2D variants is shown in Figure 7.5A; shown also for completeness in the Figure is an additional amplitude-modulated (am) ultrafast HSQC sequence previously proposed to obtain ^{1}H-^{15}N 2D protein correlations in real-time (Gal et al. (2006)). As the new sequence in Fig. 7.5A incorporates a SOFAST excitation it also possesses the main features of this experiment (Schanda et al. (2005)). Firstly, it applies all ^{1}H pulses band-selectively on the subset of protons that are of interest - in this case the amide protons - while leaving all other protons unperturbed. This assists in efficiently relaxing the excited spins, significantly decreasing their effective spin-lattice relaxation time T_1. Secondly, instead of relying on a $\pi/2$ ^{1}H excitation, a partial excitation optimized for short recycle delays (with flip angles $\alpha > 90^{\circ}$ followed by a 180° refocusing) is employed to further enhance sensitivity per unit acquisition time. The new pulse sequence also incorporates spatial encoding to monitor the indirect-domain ^{15}N-dimension using a constant-time evolution period based on bipolar gradients and linear frequency π sweeps (Pelupessy (2003)). This was preferred over alternative incremented-time versions due to the ease with which constant-time procedures can simultaneously deal with the gradients and with an effective ^{1}H decoupling. Moreover, given that we sought an indirect-domain multiple-quantum dimension free from all ^{1}H effects, two such pairs of linearly-swept ^{15}N inversion pulses were placed symmetrically with respect to the central ^{1}H refocusing pulse. In this way both the gradients as well as chemical shifts become effectively erased from the course of the ^{1}H evolution, and only ^{15}N effects are encoded.

As it was found that even with these provisions single-scan experiments were insufficient to deliver the desired SNR for the targeted protein concentration ranges, another provision was adopted to increase sensitivity: the multi-scan acquisition

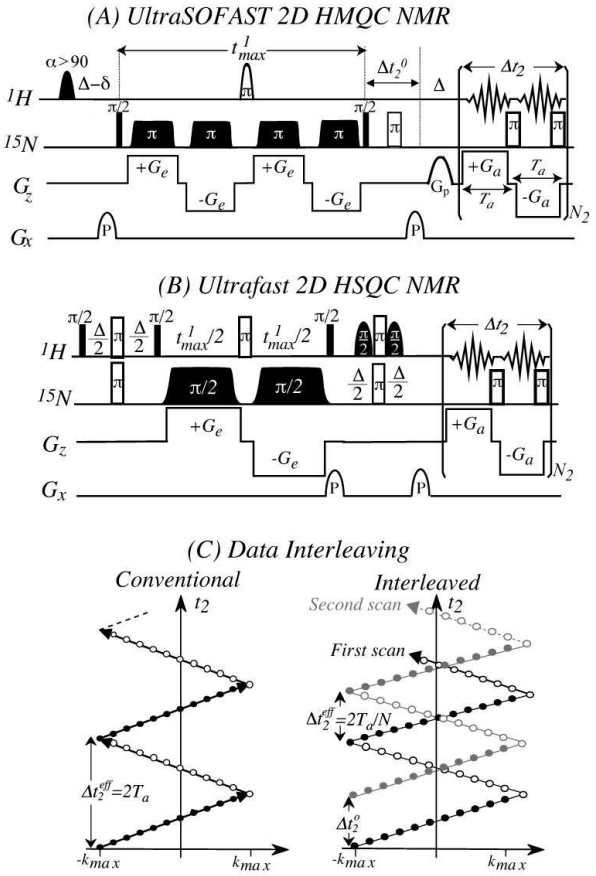


Figure 7.5 Ultrafast 2D NMR pulse sequences assayed towards the real-time characterization of protein dynamics. In all cases the frequency-chirped pulses applied for the sake of encoding the ^{15}N evolution were given WURST-like modulations and amplitudes suitable for executing $\pi/2$ or π nutations. Suppression of residual solvent signals was achieved by phase-cycling the first among the different ^{15}N pulses in concert with the receiver phase in a simple 2-step (+x,-x) cycle; further clean ups of the undesired signals were introduced by short purging gradient pulses. (A) UltraSOFAST HMQC pulse sequence introduced in this study incorporating, in addition to a constant-time spatial encoding, amide-selective PC9 ^{1}H excitation pulses (Kupce and Freeman (1994)) applied at 8 ppm with a 4 ppm bandwidth and a flip angle $\alpha > 90°$, optimized experimentally together with the subsequent REBURP (Geen and Freeman (1991)) π-pulse for a given T_{scan} time. The delay Δ was set to 5.4 ms; the delay δ accounting for spin evolution during the PC9 pulse was adjusted in a 1D version of the experiment omitting all ^{15}N pulses to yield pure-phase spectra in the ^{1}H dimension. (B) Ultrafast HSQC pulse sequence based on an amplitude-modulation of the indirect domain (Shrot et al. (2004)), incorporating Watergate (Piotto et al. (1992)), aided by the non-encoding x-gradient, for the sake of solvent suppression. (C) Principle of the data interleaving procedure, incorporated in both of the assayed sequences (A,B) for improving SNR. For the N = 2 case that is here demonstrated an initial acquisition delay Δt_2^0 is set to 0 and $\Delta t_2/2$ in alternate scans (with an intermediate ^{15}N π-pulse for J-refocusing, panel A) and data collected under $+G_a$ (filled circles) and $-G_a$ (open circles) gradients are combined in the interleaved fashion indicated by the scheme.

of interleaved data transients. Acquiring spatially-encoded scans in this manner (Fig. 7.5C) is attractive because SNR increases then with the number of interleaved scans rather than with their square root, as is traditional in signal averaging. As reminder (Matsui et al. (1985)) data interleaving is an imaging-derived procedure that introduces $\Delta t_2^0(n) = \Delta t_2(n-1)/N$ delays prior to acquisition of the $1 \leqslant$ n $\leqslant$ N transient (with $t_2 = 2T_a$ the direct-domain dwell time, cf. Fig. 7.5), and merges data obtained in such manner into a single set prior to Fourier transform along t_2. Whereas the direct-domain spectral width in single-scan acquisitions is given by $(2T_a)^{-1}$ this becomes $(2T_a/N)^{-1}$ after the data interleaving, thereby allowing one to increase the gradient oscillation time T_a by a factor N without compromising the direct-domain spectral width. This in turn allows one to reduce the acquisition gradient strength G_a by the same factor while preserving unaltered the indirect-domain spectral width, which is proportional to $k_{max} = \gamma_a T_a G_a$. As the noise penalty of ultrafast experiments is usually proportional to $G_a^{1/2}$ (due to the G_a-dependent increase in filter bandwidth) it follows that noise per scan drops by a factor of $N^{1/2}$, and the averaging of these transients leads to a full N SNR increase. Notice that in order to refocus the scalar coupling evolution occurring during these additional $\Delta t_2^0(n)$ delays a ^{15}N π-pulse was applied in their center; a two-scan $\pm 180^\circ$ phase cycling was also generally incorporated on the initial ^{15}N excitation pulse of every interleaved set, leading to $N_{scans} = 2N$.

7.2.4 Results and Discussion

Based on the 2D ultraSOFAST HMQC sequence and processing strategy just described, a series of spectra on ^{15}N-labeled Ubiquitin were acquired for a variety of sample concentrations and acquisition times. Figure 7.6 shows representative spectra recorded in this manner on 500 and 800 MHz NMR spectrometers, and compares them with a reference 500 MHz SOFAST 2D HMQC spectrum. As can be appreciated from this Figure the ultraSOFAST HMQC sequence clearly has the potential to discern a large number of resolved cross-peaks within $\sim$1 second repetition rates, for proteins at $\sim$2-3 mM concentrations. Comparisons between these 2D NMR data sets show clear resolution benefits upon going to higher magnetic fields; this is in part reflecting the relatively large line widths that all indirect-domain peaks exhibited as a result of having kept the encoding times relatively short (chosen in turn this way to prevent relaxation losses which might otherwise compromise sensitivity). On the other hand the sensitivity enhancement resulting from the increase in field is only modest, in accordance to the B_0 (as opposed to $B_0^{3/2}$) dependence expected for ultrafast experiments (Shapira et al. (2004b)).

To further explore the ultraSOFAST potential a fast kinetic process, the hy-

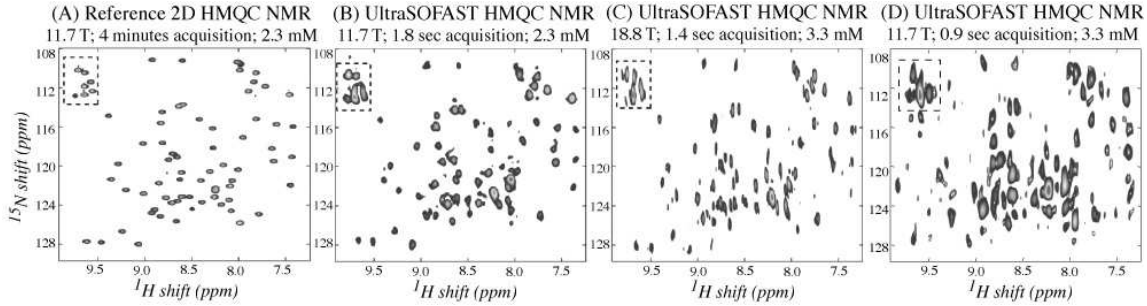


Figure 7.6 Comparison between different 2D HMQC NMR spectra of ^{15}N-labeled Ubiquitin recorded under different conditions. The analyzed solution was prepared by dissolving the His-tagged powder (Asla Biotech) in a 90%/10% H_2O/D_2O phosphate buffer (pH = 7.5). As all remaining spectra in this manuscript these data were collected at either 500 MHz on a Varian Inova console equipped with a conventional triple-axis HCN probe head, or at 800 MHz on a Bruker Avance spectrometer equipped with a QXI probe. Dashed boxes contain folded peaks arising from amine groups. (A) Reference HMQC NMR spectrum recorded at 500 MHz in 4 minutes total experimental time, serving as benchmark for our comparisons. (B) 500 MHz ultraSOFAST 2D HMQC NMR spectrum recorded using the pulse sequence in Fig. 7.5A utilizing three interleaved scans, each of them with phase-cycling of the receiver and of the ^{15}N $\pi/2$ excitation pulse. Additional experimental parameters included G_e = 40 G/cm, RF sweep range = $\pm$16 kHz, t_1^{max}= 11.6 ms, N_2 = 15, T_a = 1.43 ms, G_a = 2.7 G/cm, a 2 μs physical acquisition dwell time and $\approx$6.8 kHz analog filtering of the data. Under these conditions, indirect- and direct-domain spectral widths of $\approx$2500 and 1100 Hz were covered. Overall N_{scan} = 6 were collected at T_{scan} = 300 ms, leading to a 1.8 sec total acquisition time. Data processing involved arranging the collected FIDs in the interleaved fashion shown in Fig. 7.5C into two 715$\times$60 k/ν_1,t_2 data matrices (one corresponding to data collected with positive G_a and the other to negative G_a), zero-filling these sets to 2048$\times$128, weighting along indirect and direct dimensions with a Gaussian function, 1D FT of both data sets along t_2, magnitude calculation of the spectra, and co-addition of both data sets for the sake of increasing SNR. (C) 800 MHz ultraSOFAST HMQC spectra recorded in a total time of 1.4 seconds using four interleaved scans (N_{scan} = 8 with the phase-cycling, T_{scan} = 175 ms). Other acquisition parameters involved G_e = 31.8 G/cm, RF sweep range = $\pm$35 kHz, t_1^{max}= 14 ms, N_2 = 16, T_a = 350 μs, G_a = 12.6 G/cm, a 2 μs physical acquisition dwell time and $\approx$125 kHz analog filtering of the data. (D) Idem as (B) but involving N_{scan} = 8 with T_{scan} = 112 ms (four interleaves) resulting in a 0.9 sec total acquisition time.

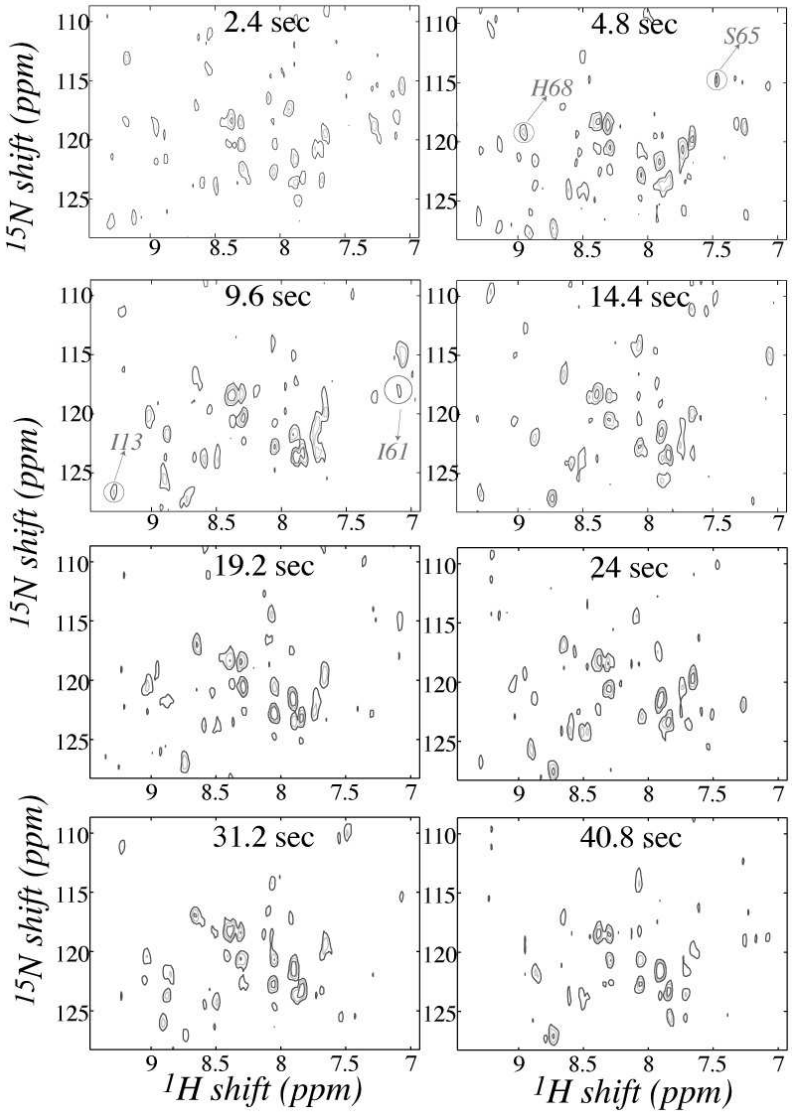


Figure 7.7 Representative series of real-time 2D ultraSOFAST HMQC NMR spectra recorded on a ≈3 mM Ubiquitin solution, following the dissolution of an initially fully protonated lyophilized powder onto a D_2O-based buffer (final uncorrected pH = 8.9). The times indicated in each frame correspond to the approximate delay elapsed since the dissolution was suddenly triggered. Acquisition parameters were akin to those in Fig. 7.6B (with N_{scan} = 8, T_{scan} = 250 ms) and so was the data processing. The repetition time between full recordings was ≈2.4 sec and data were monitored over a 20 minute interval; only a subset of the collected and processed spectra is shown. The kinetics of the highlighted peaks are depicted by the corresponding graphs in Figure 7.9.

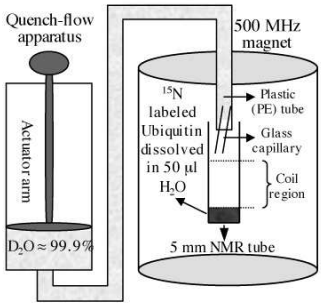


Figure 7.8 The fast mixing device used in this ultrafast H/D exchange study

drogen/deuterium exchange (H/D-exchange) in Ubiquitin was monitored in real time. Knowing such H/D-exchange rates can give valuable insight into both the thermodynamics as well as the kinetics of local and global unfolding/folding of proteins in native conditions (Arrington and Robertson (1997); Dempsey (2001); Huyghues-Despointes et al. (1999)); for the particular case of Ubiquitin such rates of H/D exchange have been recently measured by a variety of 2D techniques including ultrafast HSQC (Gal et al. (2006)) and Hadamard spectroscopy NMR (Bougault et al. (2004)).

For the present measurements an adaptation of the procedure described in ref. (Gal et al. (2006)) was used: 15 mg of ^{15}N-labeled lyophilized Ubiquitin were dissolved in 50 μl H_2O inside a Shigemi NMR tube, placed inside a properly tuned NMR system, and rapidly diluted by the addition of $\sim$400 μl of D_2O. This injection was now driven by triggering a KinTek RQF-3 quench-flow apparatus, and channeled using a small custom-made nozzle placed inside the NMR tube in order to fill the sample coil region (Scheme 7.8). Although the solvent injection lasted itself a fraction of a second, ca. 1.5 seconds had to elapse in order to let the sample turbulence settle and enable the reliable acquisition of spatially-encoded data. Even prior to this sudden D_2O injection the NMR spectrometer was already collecting ultraSOFAST HMQC spectra at a rate of one every 2.4 seconds. Further details on the parameters involved in this acquisition are given in the caption to Figure 7.7, which illustrates how the hydrogen-deuterium exchange manifests itself as a decay of the intensity of various NH resonances after the addition of D_2O. The data evidence the disappearance of various sites as a function of time, while probing life times at faster rates and shorter dead-times ($\sim$0.4 Hz and 2 sec) than had been approachable by our previous ultrafast dynamic studies for similar protein concentrations ($\sim$0.13 Hz and 30 sec respectively; Gal et al. (2006)). Some peak intensity profiles, chosen due to their relatively fast exchange kinetics, are displayed as a function of time after initiation of the experiments in Figure 7.9.

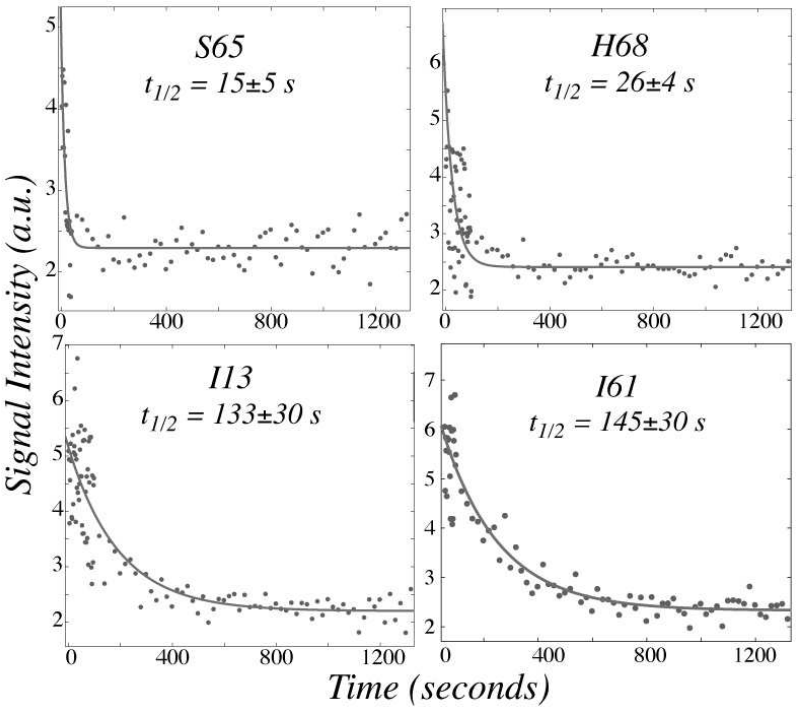


Figure 7.9 H/D exchange plots extracted from the data in Figure 7.7 for four residues exhibiting fast kinetic processes. Experimental points reflect peak heights in the ultraSOFAST HMQC 2D spectra; the $t_{1/2}$ exchange lifetimes given in the figure were obtained by fitting the data points to the equation $I(\tau) = I_o * \exp(-\tau/t_{1/2}) + I_\infty$. The points chosen over the course of the first 50 measurements (time $\tau \leqslant 120$ sec) arise from full separate 2D acquisitions, each of these 2.4 sec long. Thereafter, for the sake of reducing scattering and in view of the slowing down of the dynamics, ten consecutive spectra (50 through 60, 60 through 70, etc.) were co-added and the exchange time assigned to the center time of this average. We ascribe the shorter lifetimes arising from these fits vis-à-vis their counterparts reported in Ref. (Gal et al. (2006)), to the higher pH (8.9 vs. 6.5) in which measurements were now performed.

Given that a number of different ultrafast variants have been previously proposed and demonstrated for the recording of ^{1}H-^{15}N correlation spectra, it is pertinent to compare their general performance against the present ultraSOFAST experiments. Previous tests included HSQC versions carried out in sensitivity-enhanced and amplitude-modulated (am) modes (Pelupessy (2003); Shrot et al. (2004)). The former achieved its encoding in combination with a double reverse-INEPT element retaining both orthogonal antiphase components (Kay et al. (1992)), yielding in principle a sensitivity advantage with respect to amHSQC. In practice, however, we did not witness these theoretical expectations; perhaps owing to the for-

mer's constant-time implementation leading to an enhanced transverse relaxation, or due to its additional INEPT step susceptible to increased relaxation or pulse-imperfection losses. In any case, in view of this behavior and given our more extensive experience with amHSQC NMR, we opted for comparing this ultrafast correlation protocol against the performance of the new ultraSOFAST HMQC sequence.

With the monitoring of biomolecular dynamics as general objective, the target of such ultraSOFAST HMQC vs amHSQC comparison focused on measuring the sensitivity afforded by these different protein 2D NMR acquisition modes per unit time. Toward this end Figure 7.10A shows - using as benchmark data obtained from comparable ultrafast amHSQC and ultrafast SOFAST HMQC acquisitions - how the normalized sum of intensities arising from identical, well-resolved peaks in each spectrum changes as a function of the duration of a single scan T_{scan}. The relative ratio between the intensities afforded by these two experiments is plotted in Fig. 7.10B. These tests show that for relative long scan times (T_{scan} >2 sec) the performances of both experiments are comparable, suggesting that in this regime the potential T_2 losses of the multiple-quantum constant time ultraSOFAST HMQC evolution offset the gains that may arise from the enhanced spin-lattice relaxation. For shorter T_{scan} times, however, on the order of 100-300 ms, the sensitivity of ultraSOFAST HMQC NMR exceeds that of amHSQC by up to a factor of 3. Within the context of fast acquisitions these curves imply that if the inter-scan delay are chosen at the maximum sensitivity of the SOFAST curve (Schanda et al. (2005)), the ultraSOFAST HMQC will give sensitivity gains of ca. 2 over its amHSQC counterpart for the same total acquisition times; and that if even faster repetition rates are desired, the ultraSOFAST gain could lead to factors approaching 3 for identical experimental times.

Less satisfactory but also important to include in an SNR comparison, is the issue of water suppression. Good water suppression is of course a crucial procedure, that is challenged in rapid mixing ultrafast NMR by two non-conventional issues. One of these relates to the fact that sudden-injections of the kind illustrated in Figure 7.8 will usually take place on unshimmed samples susceptible to initial turbulence and foaming, factors which combined lead to less-than-optimal peak line widths. Another concern arises from the fact that in ultrafast NMR the spectral width associated to the physical data sampling is usually kept larger than normal to allow for the simultaneous digitization of the two spectral axes; an undesired folding-over of residuals and/or forbidden echoes from the water signal, is therefore facilitated. We found, for instance, that water-flip back schemes of the kind usually employed in standard HSQC experiments (Kay et al. (1994)) are usually insufficient for acquiring dynamic ultrafast amHSQC spectra, unless extensive phase cycling of the encoding pulses is used. UltraSOFAST HMQC, which uses se-

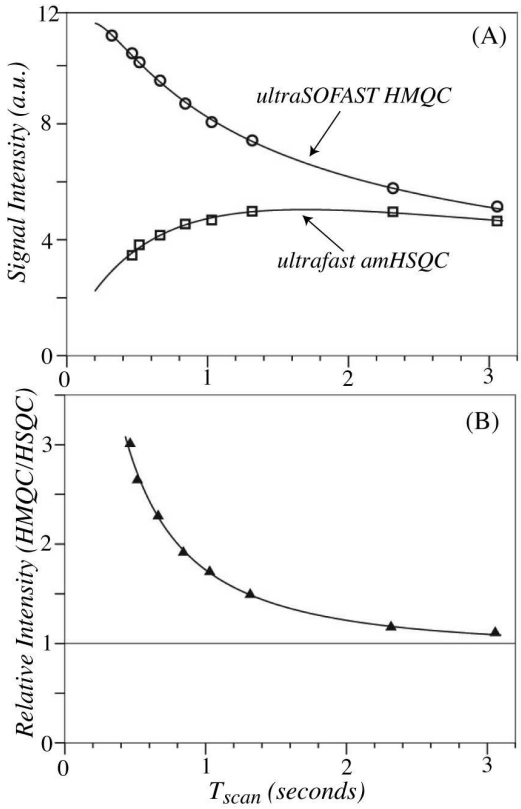


Figure 7.10 Experimental comparison between spatially-encoded ultraSOFAST HMQC and amHSQC NMR experiments, as given by the signal per unit time of 2D ^{15}N-^{1}H correlation spectra recorded with each sequence (amHSQC open triangles, ultraSOFAST HMQC open circles). The acquisition parameters were set akin to those introduced in Fig. 7.6; the duration of each frequency-chirped pulse was set to 11.6 ms for the amHSQC and 2.9 ms for the ultraSOFAST HMQC experiments, resulting in a $^{max}_{1}$ that is twice as long for the former than for the latter and thereby yielding the same effective degree of encoding for both experiments. The data points whose intensity are plotted on the top graph as a function of Tscan were obtained by summing the intensities of 17 identical, well-resolved cross-peaks in both spectra. Each spectrum was recorded in the same total time (set to 74 seconds to increase the precision of the obtained data) and scaled to the same noise level according to the number of scans. The intensity ratio between both data sets is indicated by the filled triangles of the lower plot, with a line at a ratio 1 added as visual aid.

lective pulses not to store back but rather to minimize the overall water excitation, is also handicapped in the quality of its water suppression. Indeed we found that the reliance on such selective spectral manipulations did not yield, following a basic ±x two-scan phase cycling, as complete a water suppression as that observed in ultrafast experiments incorporating Watergate-type schemes particularly when the latter relied on gradient dephasing along a transverse, non-encoding axis. The partial excitation occurring in ultraSOFAST HMQC NMR of the H_2O background was particularly noticeable in the rapid-mixing 500 MHz tests; these interferences were much weaker at 800 MHz, where the separation between the water resonance and the amide region exceeds 3 kHz. On the other hand, in experiments like those in Figure 7.7 relying on dilution by a mostly deuterated solvent, these water suppression issues were not severe. The SOFAST procedure had the added advantage of keeping the H_2O magnetization always in thermal equilibrium; this was an aid for observing amides at neutral or basic pH (e.g., pH 8.9 in Fig. 7.7) whose protons are prone to undergo rapid exchange with water molecules and thus lead to SNR losses if using saturation/dephasing of the H_2O resonance. Still, we ascribe the longer signal averaging times that were needed in the rapid mixing, mostly-deuterated solvent experiments (Fig. 7.7) over those carried out in equilibrated protonated solutions (Fig. 7.6) to the need to overcome noise problems arising from incomplete solvent suppression, rather than to inherent limitations in the available signal.

7.2.5 Conclusion

A new 2D NMR pulse scheme for the acquisition of ^{1}H-^{15}N correlation spectra at ca. 2 mM protein concentrations and at repetition rates approaching 1 Hz was presented, combining the benefits of SOFAST HMQC and of spatial encoding. Although previously presented spatially encoded techniques had been able to yield a 2D protein HSQC NMR spectra in a single scan, practical sensitivity considerations makes this possible only at relatively high concentrations (ca. 3 mM for state-of-the-art hardware). Moreover, monitoring protein dynamics by executing multiple such HSQC acquisitions at high repetition rates, would still be challenged by the usual requirements of spin relaxation. The proposed combination of SOFAST and ultrafast techniques is definitely a valuable step towards alleviating this situation, enabling the repetition of 2D NMR measurements at the highest possible frame rate. In fact SOFAST HMQC NMR has already been shown compatible with another fast acquisition method, Hadamard-encoding, a combination which decreases the number of necessary scans significantly and also allows one to record 2D ^{1}H-^{15}N correlations within second-long timescales (Schanda and Brutscher (2006), see section 7.1). Unlike Hadamard spectroscopy, however, the pro-

posed spatially-encoded ultraSOFAST HMQC experiment results in full 2D spectral correlations rather than in a series of 1D projections, yielding an unbiased range of spectral changes along both dimensions. We believe that this may be an aid to study processes where peak positions might be changing due to the dynamics, as would be the case upon following protein folding. We trust that, with further increases in sensitivity brought about by advances in magnet and probe technology and with additional pulse-sequence developments including better approaches to solvent suppression, the proposed experiment will eventually become a useful route to study fast biomolecular kinetics at a site-resolved level and within sub-second timescales.

8 HET-SOFAST NMR: Longitudinal proton relaxation exploited for fast detection of structural compactness and heterogeneity along polypeptide chains

8.1 Abstract

Structure elucidation of proteins by either NMR or X-ray crystallography often requires the screening of a large number of samples for promising protein constructs and optimal solution conditions. For large-scale screening of protein samples in solution, robust methods are needed that allow a rapid assessment of the degree of structural compactness of a polypeptide under diverse sample conditions. Here we present HET-SOFAST NMR, a highly sensitive new method for semi-quantitative characterization of the structural compactness and heterogeneity of polypeptide chains in solution. Based on one-dimensional ^{1}H HET-SOFAST NMR data, obtained on well-folded, molten globular, partially- and completely unfolded proteins, we define empirical thresholds that can be used as quantitative benchmarks for protein compactness. For ^{15}N-enriched protein samples, two-dimensional ^{1}H-^{15}N HET-SOFAST correlation spectra provide site-specific information about the structural heterogeneity along the polypeptide chain.

8.2 Introduction

A large amount (30 to 60%) of the protein sequences encoded in the eukaryotic genomes is predicted to be predominantly unstructured. Dunker and Obradovic

(2001); Uversky (2002) In structural proteomics projects, concerned with the three-dimensional structure determination of all proteins from a particular genome, a significant fraction of the expressed proteins is thus expected to be at least partially unfolded. For these proteins, structure elucidation by either NMR or X-ray crystallography will be difficult, if not impossible, in the absence of an appropriate binding partner that increases the structural compactness. Therefore, experimental methods are required that allow fast identification of promising protein constructs and sample conditions for structure determination. Besides being a well-established technique for obtaining structural and dynamic information at atomic resolution of moderately sized proteins in solution, NMR is also a powerful tool for fast screening of protein samples for the presence of stable structures along the polypeptide chain. (Prestegard et al. (2001); Rehm et al. (2002); Staunton et al. (2003)) A close look at the 1D ^{1}H NMR spectrum provides a first idea of the structural compactness and heterogeneity of the polypeptide chain(s) present in the NMR sample tube. The ^{1}H chemical shift dispersion is indicative of the presence of secondary and tertiary structure. Konrat and coworkers have shown recently that a statistical analysis of the ^{1}H spectrum yields a global measure of the structural compactness of the protein(Hoffmann et al. (2005)). Furthermore, the line shape of the peaks reports on the molecular dynamics, and especially on the rotational molecular tumbling correlation time that is directly related to the oligomerization state of the protein. If ^{15}N isotope enrichment is available, a two-dimensional ^{1}H-^{15}N correlation spectrum provides further information on the purity of the sample by comparing the number of detected cross peaks with the number of non-proline residues in the polypeptide sequence. Large differences in peak intensities among the correlation peaks in the 2D spectrum are indicative of structural and dynamic heterogeneities. Other more sophisticated NMR experiments provide semi-quantitative information on the protein structure and dynamics. One of the most prominent examples is the measurement of the ^{1}H-^{15}N heteronuclear NOE that can be related to the amplitude of fast local motions experienced by the N-H vectors along the polypeptide chain. The measurement of amide-water hydrogen exchange rates (Hwang et al. (1997); Krishna et al. (2004)) or ^{1}H-^{1}H spin diffusion (noe) effects (Kutyshenko and Cortijo (2000)) has also been proposed in the past as a measure of structural compactness and rigidity in native and molten globular states of proteins. Here we present new 1D ^{1}H and 2D ^{1}H-^{15}N NMR experiments that measure the effect on the amide ^{1}H spectrum of either perturbing the water ^{1}H or aliphatic ^{1}H spin polarization. The 1D experiments yield global observables λ_{ex} and λ_{noe} that provide a measure of the overall solvent accessibility (λ_{ex}) and proton density (λ_{noe}) at the amide proton sites, and thus the two parameters yield complementary information on the structural compactness of the protein. In the 2D version, λ_{ex} and λ_{noe} values are obtained for each resolved correlation peak, pro-

viding a means of assessing the structural and dynamic heterogeneity along the polypeptide chain. The experiments are simple and easily implemented on any standard NMR spectrometer. The high sensitivity of the pulse schemes allows application to low-concentrated protein samples with acquisition times ranging from a few seconds to a few minutes. Fast 1D and 2D data acquistion is particularly attractive for large-scale screening of proteins under various sample conditions (pH, temperature, salt concentration, osmolytes, etc.), or in the presence of different ligands.

8.3 HET-SOFAST NMR experiments

In order to measure the degree of structure in a particular protein or protein fragment, we have developed one-dimensional (1D) and two-dimensional (2D) versions of what we like to call a HETerogeneity-SOFAST experiment. The pulse sequences of the 2D ^{1}H-^{15}N HET-SOFAST-HMQC and 1D ^{1}H HET-SOFAST experiments are displayed in Figs. 8.1a and b, respectively. The pulse schemes are derived from a band-Selective Optimized-Flip-Angle Short-Transient (SOFAST)-HMQC experiment. SOFAST-HMQC combines the advantages of a small number of radio-frequency pulses, Ernst-angle excitation, and longitudinal relaxation optimization to obtain significantly increased sensitivity for high repetition rates of the pulse sequence. In the new HET-SOFAST NMR experiments an additional band-selective inversion pulse is added at the beginning of the recycle period or contact time t_c on the H_{sat} proton channel. The experiment measures the effect of this perturbation (saturation) on the spin polarization of the amide protons after a contact time t_c. Two data sets are recorded, with and without the H_{sat} inversion pulse. An example of such a set of 2D ^{1}H-^{15}N HET-SOFAST-HMQC spectra recorded for a 2.0 mM sample of native-state ubiquitin (600 MHz, 25°C) in only 10 seconds of acquisition time per data set is shown in Figs. 8.1c and d, illustrating the high sensitivity obtained by these experiments. The ratio of the peak intensities ($\lambda = I_{sat}/I_{ref}$) measured in the two spectra then provides insight into the interaction between amide protons (H^N) and saturated protons (H^{sat}). For the 1D version of the HET-SOFAST experiments, shown in Fig. 8.1b, the ^{15}N pulses and the incremented time period (t_1) are simply omitted.

Two types of proton interactions may be monitored by the HET-SOFAST experiments of Fig. 8.1. If the H_{sat} pulses are applied to the aliphatic protons, the experiment measures the ^{1}H-^{1}H spin diffusion or NOE effect (λ_{noe}). If the water polarization is inverted, the intensity ratios (λ_{noe}) provide a measure of the amide-water hydrogen exchange rates. The two observables, λ_{noe} and λ_{ex} yield complementary information on the structure and mobility of the polypeptide chain. The

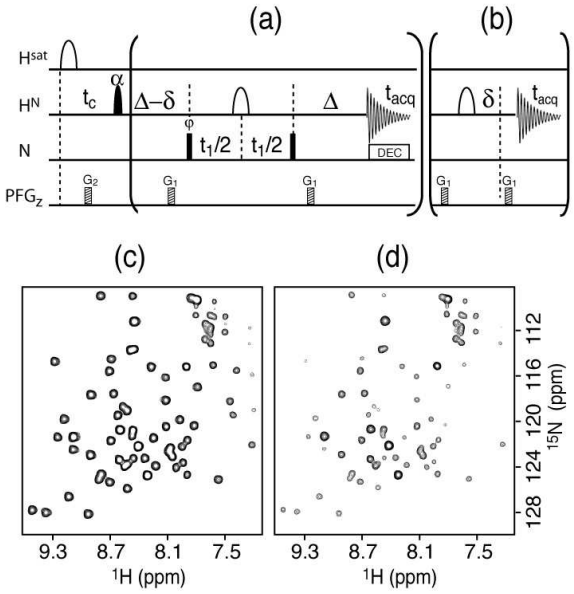


Figure 8.1 Pulse sequences to record (a) 2D ^{1}H-^{15}N HET-SOFAST-HMQC and (b) 1D ^{1}H HET-SOFAST NMR spectra. The pulses and gradient outside the brackets are common to both 1D and 2D experiments. Filled and open pulse symbols indicate 90° and 180° rf pulses, except for the H^{N} excitation pulse applied with flip angle α (typically set to 120°). The variable-flip-angle HN pulse has a polychromatic PC9 shape and band-selective HN refocusing is realized using REBURP (Geen and Freeman (1991)). The HN pulses are centered at 9.5 ppm and cover a band width of 4.0 ppm. For the HET$_{ex}$-SOFAST experiments a selective H^{sat} inversion pulse is applied at the water frequency with a Gaussian shape and a pulse duration of 100ms at 600 MHz, corresponding to a band-width of 9.0 Hz. For aliphatic ^{1}H inversion (H^{sat}) in HET$_{noe}$-SOFAST experiments an I-SNOB5 (Kupce et al. (1995)) pulse is applied 3.0 ppm shifted from the water resonance, and covering a band width of 4.0 ppm, corresponding to a pulse length of 1.88 ms at 600 MHz. The transfer delay Δ is set to $1/(2J_{HN}) \approx 5.4$ ms, and t_c is the inter-scan delay or contact time. The delay δ (≈1.6 ms) accounts for spin evolution during the PC9 pulse, and has to be adjusted prior to data acquisition to yield pure-phase spectra in the ^{1}H dimension. Pulsed field gradient pulses G1 and G2 are applied along the z-axis (PFGz) for solvent and artefact suppression. They are typically applied with gradient strength of 10-20 G/cm during 200μs followed by a recovery delay of 100 μs. To avoid extensive first order phase corrections in the t_1 dimension of experiment (a), the 90°-t_1-90° pulse sequence element on the ^{15}N channel is applied simultaneously to the REBURP pulse on the ^{1}H channel. Quadrature detection in t_1 is obtained by phase incrementation of φ according to STATES or STATES-TPPI. Adiabatic WURST-2 ^{15}N decoupling (Kupce and Wagner (1995)) is applied at a field strength of $\gamma B_1/2\pi = 550$ Hz. Additional HETnoe-SOFAST-HMQC spectra recorded without (ref-spectrum) and with (sat-spectrum) aliphatic spin inversion are plotted in (c) and (d), respectively. The 2D data were recorded at 600 MHz ^{1}H frequency on a sample of 2.0 mM ^{15}N-labeled ubiquitin using the following acquisition parameters: α=120, Δ=5.4 ms, δ = 1.57 ms, t_{acq}^{max}=40 ms, t_c=50 ms. 40 complex points were recorded for t_1^{max}=29 ms, resulting in a total experimental time of 10 s per spectrum. Site-specific λ_{noe} values can be extracted from the spectra in (c) and (d) by calculating the intensity ratio I_{sat}/I_{ref} for each cross peak.

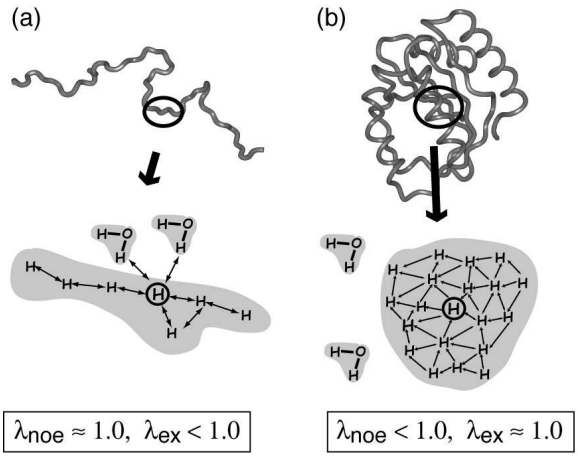


Figure 8.2 Illustration of ^{1}H-^{1}H spin diffusion and amide-water hydrogen exchange effects in (a) an elongated polypeptide chain, characterized by a low-density flexible ^{1}H network and a high water accessibility of the amides, and (b) a globular well-structured protein characterized by a high density rigid ^{1}H network and a low water accessibility of the amides.

λ_{noe} values report on the local ^{1}H spin network, i.e. the number, density, and mobility of protons close to a particular amide, whereas the λ_{ex} values are indicative of the water accessibility of the amide hydrogens. As illustrated in Fig. 8.2, a hydrogen atom inside a well structured globular protein is surrounded by many other hydrogens, but generally not easily accessible to water molecules (Fig. 8.2b). This situation allows efficient proton spin diffusion ($\lambda_{noe} << 1$), but only little hydrogen exchange ($\lambda_{ex} \approx 1$). The inverse is found in the case of an elongated peptide chain as depicted in Fig. 8.2a, characterized by a low proton density ($\lambda_{noe} \approx 1$) and a high solvent accessibility ($\lambda_{ex} << 1$).

8.3.0.1 HET$_{noe}$-SOFAST: measurement of ^{1}H-^{1}H spin diffusion or NOE

The dipolar coupled spin network in a folded polypeptide chain provides the main mechanism for spin-lattice relaxation of the ^{1}H spins after a perturbation of the equilibrium spin polarization. If only a few ^{1}H spins, e.g. all amide ^{1}H, are excited by means of selective rf pulses, the energy put into the spin system is transferred efficiently within the ^{1}H spin network via ^{1}H-^{1}H NOEs, and evacuated to the lattice at the sites of high mobility, e.g. methyl groups or highly-flexible parts of the peptide chain (selective case). This is the situation found during the contact time t_c in the reference experiment. In the saturation experiment, where most of the ^{1}H spins are out of their thermal equilibrium state at the beginning of t_c, the ^{1}H-^{1}H spin diffusion efficiency is considerably reduced, resulting in less efficient energy

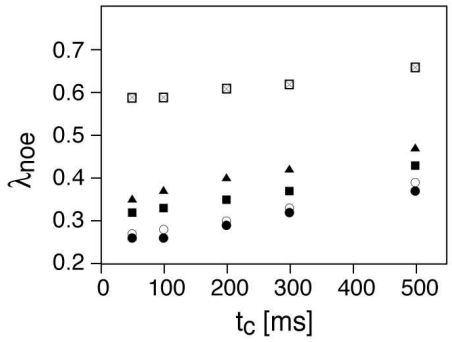


Figure 8.3 Experimental evaluation of the influence of the contact time t_c and the sample temperature (tumbling correlation time) on the measured λ_{noe} values. 1D HET$_{noe}$-SOFAST NMR spectra were recorded at 600 MHz on a sample of cytochrome c553 (9 kDa), and the measured intensity ratios are plotted as a function of the contact time for data measured at 15°C ($\tau_c \approx 6.6$ ns, filled circles), 25°C ($\tau_c \approx 5.0$ ns, open circles), 35°C ($\tau_c \approx 3.9$ ns, filled squares), and 45°C ($\tau_c \approx 3.1$ ns, filled triangles). In addition, open squares represent λ_{noe} values measured for urea-denatured ubiquitin at 25°C.

dissipation in the dipolar-coupled ^{1}H spin network (non-selective case). HET$_{noe}$-SOFAST NMR provides a measure of the ratio of effective spin-lattice relaxation rates in the selective (reference) and non-selective (saturation) experiments.

The ^{1}H-^{1}H spin diffusion effect, and thus the measured λ_{noe} values depend on the effective tumbling correlation time (τ_c) and the contact time (t_c). To evaluate experimentally the effect of these two parameters on the measured intensity ratios, we have performed 1D HET$_{noe}$-SOFAST experiments on a sample of cytochrome c553 at 600 MHz ^{1}H frequency using different contact times (50 to 500 ms). In order to estimate the influence of the molecular tumbling correlation time, the experiments were repeated for different sample temperatures varying from 15°C to 45°C. In addition, we have measured λ_{noe} values as a function of the contact time t_c for a sample of urea-denatured ubiquitin. Several conclusions can be drawn from the results of this comparison, shown in Fig. 8.3. First, one observes an increase in the λ_{noe} values for increasing contact times as expected for a cross-relaxation effect. Interestingly this increase is less pronounced for the unfolded polypeptide chain than for the globular protein. Therefore the use of short contact times increases the dynamic range of structural compactness that can be probed by the λ_{noe} values. A second observation is that the spread in λ_{noe} values obtained for different temperatures (mainly altering the overall tumbling correlation times) of the folded peptide chain is quite independent of the contact time used. The covered range on the λ_{noe} scale is about 0.1 for τ_c values ranging from about 3 to 7 ns. A part of these λ_{noe} variations is expected to be a consequence of the reduced fold stability at higher

temperatures. The observed range of about 0.1 can thus be considered an upper limit of the effect of the tumbling correlation time on the measured λ_{noe} values. Note that for slower tumbling molecules ($\tau_c > 7$ ns) we do not expect a significant further decrease in λ_{noe}, as the τ_c-dependent ^{1}H-^{1}H spin diffusion process will no longer be the limiting step for the ^{1}H spin-lattice relaxation efficiency in the reference experiment. For slowly tumbling globular proteins, the λ_{noe} values are mainly determined by the internal flexibility of the polypetide chain, responsible for the energy dissipation to the lattice. In practice, λ_{noe} values below 0.2 are rather unlikely to be detected for the contact times of around 100 ms used in this study (see Figure 8.5). We have also experimentally tested the influence of the HN excitation flip angle (α) on the measured λ_{noe} ratios. No significant changes were observed when changing the excitation angle between 90° and 120° (data not shown). Therefore, for short contact times in the range 50 to 200 ms, a flip angle $\alpha \approx 120°$ will yield optimal sensitivity. (Schanda and Brutscher (2005); Schanda et al. (2005)) In addition, no changes in the measured λ_{noe} values are expected from variations in the pH of the solution. Therefore, 1D HET$_{noe}$-SOFAST measurements on different protein samples using a fixed contact time allows to compare the behavior of these proteins in terms of structural compactness, independent of their molecular weight and the particular sample conditions used.

8.3.0.2 HET$_{ex}$-SOFAST: measurement of amide-water hydrogen exchange

Hydrogen atoms at the polar groups of the polypeptide chain such as amides are in continual exchange with solvent hydrogens. The exchange rates critically depend on the pH, the local pK$_a$ values, the temperature, and the solvent accessibility of the individual amides. As a consequence of the differential protection (H-bonding) in a folded polypetide chain, the characteristic exchange times are spread over several orders of magnitude (milliseconds to days or months). The measurement of hydrogen exchange rates for amides has been used for many years in protein NMR to probe differential solvent protection along the polypeptide chain. (Krishna et al. (2004); Wagner and Wuthrich (1982)) The HET$_{ex}$-SOFAST experiment of Fig. 8.1 has been designed to monitor fast hydrogen exchange. The experiment provides a measure of the ^{1}H spin polarization that is transferred between the exchangeable amide and water hydrogen atoms after selective perturbation of the water resonance. In order to detect a measurable effect, a significant fraction of amide hydrogen atoms must have been exchanged with solvent hydrogens during the contact time t_c. For contact times of a few hundred milliseconds, this means that the HET$_{ex}$-SOFAST experiment is sensitive to hydrogen exchange at the sub-second time scale. The experiment is thus most appropriate for samples at neutral or basic pH and ambient temperature, where hydrogen residence times of less than a sec-

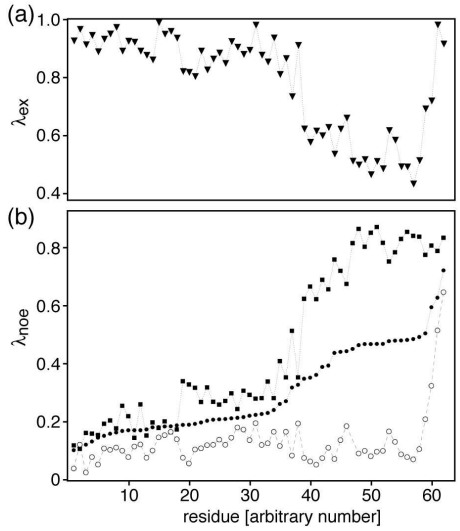


Figure 8.4 Influence of H_{sat} pulse shape parameters on the measured λ_{noe} values. (a) λ_{ex} and (b) λ_{noe} values measured for a heterogeneous polypeptide chain at neutral pH. In all experiments the ^{1}H carrier was positioned on the water resonance (4.7 ppm). For the HET$_{noe}$-SOFAST measurements, the aliphatic ^{1}H inversion pulse (H_{sat}) was centered at 1.7 ppm covering a band width of 4.0 ppm (filled squares), at 2.0 ppm covering a band width of 4.0 ppm (filled circles), and at 2.0 ppm covering a band width of 6.0 ppm (open circles). The residues (correlation peaks) have been sorted for increasing λ_{noe} values (filled circles). The different pulse calibrations mainly affected the measured λ_{noe} ratios for amides characterized by low λ_{ex} values, while the noe ratios for amides with high ex values were only little changed.

ond are observed for amides located in the flexible, loosely structured parts of the polypeptide chain. The strong dependency on solvent conditions, however, makes it difficult to compare the λ_{ex} values measured for different protein samples.

8.3.0.3 Optimization of shaped ^{1}H pulse parameters

The HET-SOFAST experiments require selective spin manipulations of amide, aliphatic, and water protons. It is therefore important that the ^{1}H pulse shapes are adjusted in a way that these proton species can be considered as spectroscopically independent (similar to heteronuclear spin systems). The amide and aliphatic ^{1}H pulses have been adjusted to yield minimal perturbation of the water resonance, while covering most of the amide or aliphatic ^{1}H spectral region. For the inversion of the water resonance a compromise between pulse duration and selectivity is required. The optimized pulse parameters are given in the caption to Fig. 1. To demonstrate the importance of pulse parameter optimization for the measurement of λ_{noe} values, the results of an experimental comparison using different aliphatic

^{1}H inversion pulses are shown in Fig. 8.4. If the aliphatic pulses affect the water spin polarization (filled and open circles in Fig. 8.4b), the measured λ values are a mixture of spin diffusion (λ_{noe}) and exchange effects (λ_{ex}). As a consequence of this interference, low λ_{noe} values will be measured for the well-structured part of the polypeptide chain because of the ^{1}H-^{1}H spin diffusion effects, while hydrogen exchange effects with the (partially) saturated water will also yield reduced λ_{noe} values for the flexible water-accessible sites. This partly hides the information on the structural (and dynamical) heterogeneity along the polypeptide chain. Fig. 8.4 demonstrates that a careful choice of the selective pulse parameters renders the two effects independent.

8.4 Application to protein sample screening and characterization

1D and 2D HET-SOFAST experiments were performed on a Varian INOVA 600 spectrometer equipped with a triple resonance (room-temperature) probe and shielded z-gradients. Unless indicated in the text, data were acquired at a sample temperature of 25°C. Data processing was achieved using either FELIX (Biosym Technologies) or NMRPipe(Delaglio et al. (1995)) Typically, the 1D data were multiplied by a squared cosine function and zero-filled to 8096 complex data points prior to Fourier transformation. Intensities I_{sat} and I_{ref} were obtained by integrating the spectrum from about 7.0 to 12.0 ppm. 2D data sets were zero-filled to final matrices of 2048 (^{1}H) $\times$ 1024 (^{15}N) complex data points. Spectral resolution was increased by signal apodization in the ^{1}H dimension using a 40° shifted cosine function, and by forward linear prediction in the ^{15}N dimension. The maximum peak heights obtained from a grid search algorithm were used as I_{sat} and I_{ref} values for the calculation of the amide-specific $\lambda_{noe/ex}$ values.

The following protein samples were used in this study: human ubiquitin (8.6 kDa) native (N)-state (Vijay-Kumar et al. (1987)) at pH 6.2, 2.0 mM, ubiquitin A-state (Brutscher et al. (1997)) in 60%/40% methanol/water mixture at pH 2.0, 0.7 mM, ubiquitin denatured in 8M urea (U-state) at pH 2.0, 1.0 mM (Peti et al. (2001)), D. vulgaris cytochrome c553 (Blackledge et al. (1995)) (9.0 kDa) at pH 5.9, 5.0 mM, human α-synuclein (14.5 kDa) in SDS micelles (Ulmer and Bax (2005)) at pH 7.0, 0.5 mM, bovine α-lactalbumin (16.2 kDa) at pH 7.2, 0.6 mM (Wijesinha-Bettoni et al. (2001)), bovine α-lactalbumin in molten globular state (MG) at pH 2.5 (Arai and Kuwajima (2000)), 1.0 mM,S. hindustanus bleomycine-resistance protein (BRP, 30.4 kDa) in free form and in complex with Zn^{2+}-ligated bleomycine (Vanbelle et al. (2003)) at pH 6.5, 1.0 mM, E. coli SiR-FP18 (Champier et al. (2002)) (18.0 kDa) at pH

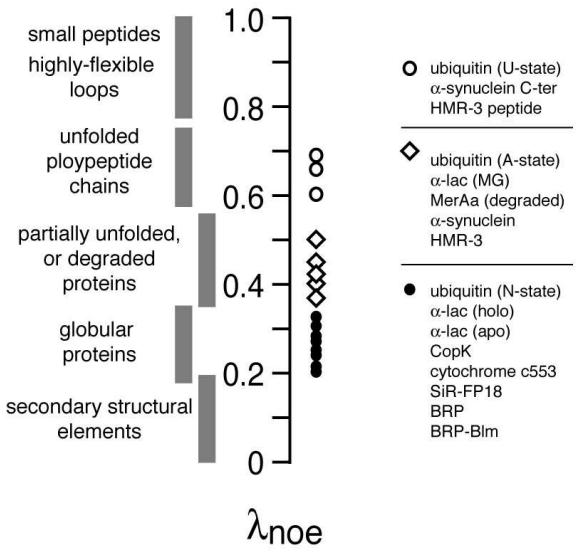


Figure 8.5 Classification of polypeptide samples, based on measured λ_{noe} values, in categories corresponding to different degrees of structural compactness. The classification is based on thresholds determined form experimental 1D HETnoe-SOFAST NMR data measured on various protein samples at 600 MHz using a contact time of t_c=100 ms. For accurate and reproducible results it is important to fine-tune the power level of the REBURP HN refocusing pulse for minimal water and maximal amide proton signal in a single-scan 1D reference SOFAST experiment.

7.0, 1.5 mM, and 3 proteins involved in heavy-metal resistance in R. metallidurans CH 34 (Mergeay et al. (2003)), MerAa (Rossy et al. (2004)) (8.2 kDa) at pH 7.5, 1.5 mM, CopK (Tricot et al. (2005)) (8.3 kDa) at pH 6.5, 1.5 mM, and a third protein with a molecular weight of 13.2 kDa which we will refer to as HMR-3 (for confidential reasons) at pH 6.0, 0.9 mM.

8.4.0.4 1D HET-SOFAST NMR: semi-quantitative characterization of the overall compactness of polypeptide chains

1D ^{1}H NMR is a valuable tool for fast NMR screening of protein samples as it does not require isotope enrichment, and because spectra can be acquired in a very short experimental time (typically a few minutes) for samples as dilute as several μM on a high field NMR spectrometer equipped with a cryogenic probe. Although the 1D ^{1}H spectrum contains a great amount of information on the structure and dynamics of the polypeptide chain(s) in the NMR sample, it is not straightforward to quantify this information. Here we show that 1D HET$_{noe}$-SOFAST NMR can be used as a very simple tool for classification of the overall compactness of a particular polypeptide chain. Fig. 8.5 shows the results of 1D HET$_{noe}$-SOFAST measure-

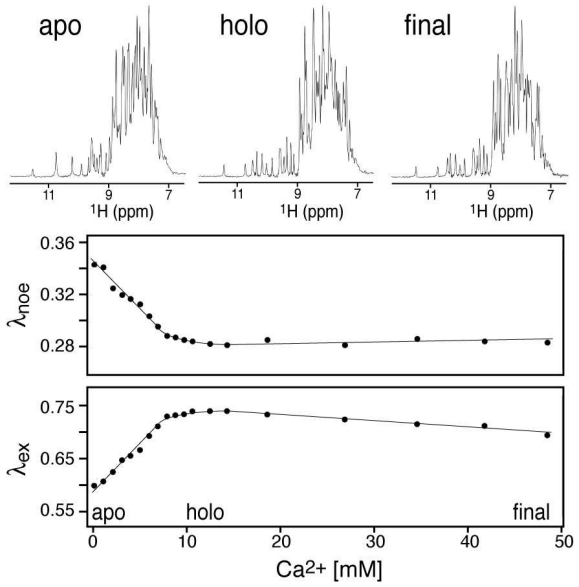


Figure 8.6 Calcium-induced changes in structural compactness of the protein α-lactalbumin monitored by 1D HET-SOFAST NMR at 600 MHz using a contact time of t_c=200 ms. The measured λ_{noe} and λ_{ex} values are plotted as a function of the Ca^{2+} concentration. In addition, ^{1}H spectra (amide region) are plotted for the different protein states: calcium-free form (apo), calcium-bound form (holo) and calcium-saturated form (final).

ments for a representative number of protein samples available in our laboratory. These samples vary in structure, molecular weight, and solvent conditions. The measured λ_{noe} values can be classified into 3 categories: (i) λ_{noe} = 0.2 to 0.35: well-folded globular proteins with a limited number of loop regions; (ii) λ_{noe} = 0.35 to 0.55: polypeptide chains with large parts of unstructured peptide fragments, or protein samples containing a significant amount of small peptide chains (e.g. from protein degradation); (ii) λ_{noe} = 0.55 to 0.75: mostly unfolded (random coil) polypeptide chains. λ_{noe} values below 0.2 or above 0.8 are generally not found for larger polypeptide chains, but they are indicative of compact secondary structural elements and small flexible peptide chains, respectively. Such λ_{noe} values are observed in 2D ^{1}H-^{15}N HET$_{noe}$-SOFAST spectra (see below).

We have also investigated the possibility of using 1D HETnoe-SOFAST and 1D HET$_{ex}$-SOFAST NMR to monitor subtle changes in the overall compactness of a polypeptide chain upon variation of solvent conditions. Bovine α-lactalbumin, the regulatory component of the lactose synthase complex, presents an interesting test case for such a study. In the apo form, α-lactalbumin forms a compact tertiary structure comprising 4 α-helices, several short 310 helices, and a triple-stranded antiparallel β-sheet (Pike et al. (1996)). α-lactalbumin has a conserved

high-affinity Ca^{2+} binding site, and it has been shown that upon Ca^{2+} binding (holo form) the protein shows an increased stability but only little change in the overall structure (Chrysina et al. (2000); Wijesinha-Bettoni et al. (2001)). It is therefore interesting to test whether the HET-SOFAST experiments are able to monitor these slight changes in the overall compactness of α-lactalbumin upon Ca^{2+} binding. The results of a titration study are shown in Fig. 8.6. The Ca^{2+}-depleted apo form of α-lactalbumin was obtained by adding 10 mM EDTA to a ca. 0.6 mM protein solution (pH 7.3). The Ca^{2+} titration was achieved by adding small quantities of a concentrated stock solution of $CaCl_2$ (same buffer and pH) to the apo-sample. We observe a decrease in the measured λ_{noe} values from 0.34 to 0.28, and an increase in the λ_{ex} values from 0.60 to 0.75 when adding CaCl2. These observations are in agreement with a polypeptide chain becoming more and more compact with increasing Ca^{2+} concentration. A maximum (minimum) of the titration curves is reached at a Ca^{2+} concentration of about 10-12 mM, where both EDTA and the high-affinity binding site of α-lactalbumin are occupied by a Ca^{2+} ion. For even higher calcium levels the λ_{noe} (and λ_{ex}) value again slightly increases (and decreases). This finding is explained by the presence of a second Ca^{2+} binding site of lower affinity. It has been shown by analytical affinity chromatography that the occupancy of this second binding site increases the surface hydrophobicity, suggesting a destabilization of the protein structure similar to the apo form.(Fitzgerald and Swaisgood (1989))

The availability of both observables λ_{noe} and λ_{ex} makes sure that the observed changes are not due to slight differences in the tumbling correlation time or in the pH of the sample, but they are explained by a change in the overall structural compactness. The experimental results obtained for α-lactalbumin clearly demonstrate that HET-SOFAST NMR is able to detect subtle changes in the structural compactness and stability of polypeptide chains in different solution conditions. Monitoring such slight changes is certainly not easily accomplished by a simple inspection of the corresponding 1D spectra (see Fig. 8.6).

8.4.0.5 2D HET-SOFAST NMR: fast characterization of the structural heterogeneity of polypeptide chains

Site-specific structural information in protein NMR requires the recording of (at least) 2D correlation spectra. In the context of fast sample screening, ^{1}H-^{15}N experiments are especially attractive because of their high intrinsic sensitivity that enables short acquisition times for moderately concentrated samples. As the required ^{15}N isotope enrichment is relatively inexpensive as long as the protein is obtained from a suitable expression system, it is quite feasible to acquire such data for a large number of samples during the screening of optimal sample conditions or in

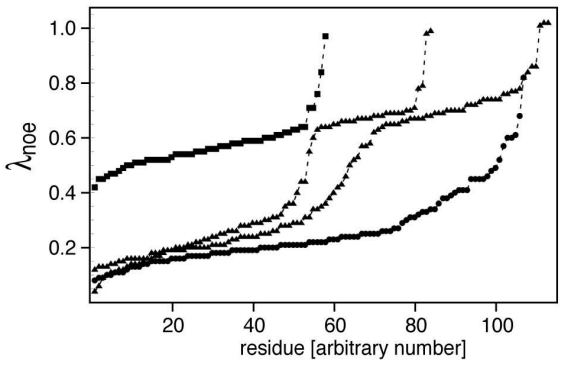


Figure 8.7 λ_{noe}-dispersion profiles obtained for several protein samples from a set of 2D HET-SOFAST-HMQC spectra recorded at 600 MHz using t_c=200 ms: a well-structured globular protein (circles), two heterogeneous proteins comprising structured and unstructured domains (triangles), and a polypeptide chain under denaturing conditions (squares). The residues (correlation peaks) for all 4 proteins have been sorted for increasing λ_{noe} values.

a high throughput sample preparation pipeline. The 2D HET-SOFAST-HMQC experiments introduced in Fig. 8.1 present a simple and fast NMR method for measuring the amount of local structure and dynamics along the polypeptide chain on a residue-by-residue basis. Fig. 8.7 shows the λ_{noe} profiles obtained with the 2D HET$_{noe}$-SOFAST-HMQC pulse sequence for 4 different protein samples, representative for a well-folded globular protein (BRP), an unfolded polypeptide chain (ubiquitin U-state), and 2 heterogeneous proteins (α-synuclein, HMR-3) comprising both, folded and unfolded domains. Note that a third example of a heterogeneous peptide chain is shown in Fig. 8.3. The measured intensity ratios have been sorted by increasing λ_{noe} value. Quite distinct profiles are obtained for the different types of proteins. For the globular protein the majority of residues show λ_{noe} values below 0.3 indicative of well-structured compact peptide segments. Higher values are obtained for the loop regions connecting these secondary structural elements. For the unfolded peptide chain most λ_{noe} values are between 0.4 and 0.6, indicating the absence of any permanent structure in this particular polypeptide chain. Finally, the two heterogeneous protein constructs display the features of both, structured and unstructured proteins. A significant number of residues have λ_{noe} values below 0.3, corresponding to a well-structured protein fragment. But a second "plateau" is detected around λ_{noe} = 0.6, indicative of a random coil structure of the corresponding protein fragment. This illustrates that, in the absence of any resonance assignment, 2D HET$_{noe}$-SOFAST-HMQC can be used in a fast and straightforward way for the characterization of the structural heterogeneity, and

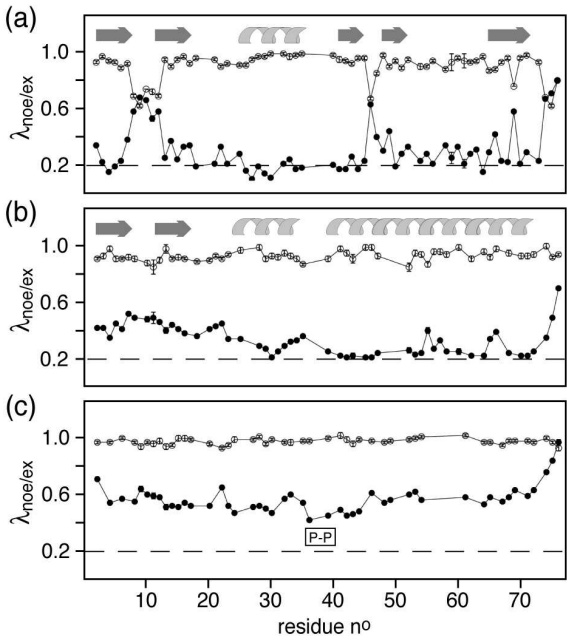


Figure 8.8 2D HET$_{noe}$-SOFAST-HMQC and HET$_{ex}$-SOFAST-HMQC experiments performed on (a) the native N-state, (b) the methanol-induced A-state, and (c) the urea-denatured U-state of ubiquitin. The measured λ_{noe} (filled circles) and λ_{ex} (open circles) intensity ratios are plotted as a function of the polypeptide sequence. The dashed horizontal line corresponding to λ_{noe}=0.2 has been added as a visual aid for the comparison of the individual data sets. Structural elements are indicated on top of (b) and (c), whereas the proline motive P37-P38 is highlighted in (a).

for the estimation of the relative percentage of folded and unfolded parts of the polypeptide chain.

Once the backbone resonance assignment is known for a protein sample, the same 2D HET$_{noe}$-SOFAST-HMQC data can be used to identify or confirm the presence and absence of local structure along the polypeptide chain. This is shown in Fig. 8.8 for the polypeptide chain of ubiquitin in several structural states obtained using different solvent conditions: (a) native (N) state (water, pH 6.2), (b) A-state (60% methanol, 40% water, pH 2.0), and (c) U-state (8M urea, pH 2.0). For the N-state, only some loop regions appear to be less compactly structured than the rest of the protein with the lowest λ_{noe} values observed for the central part of the α-helix. The A-state also displays a quite homogeneous behavior with λ_{noe} values ranging from 0.2 to 0.5. This confirms that the A-state of ubiquitin is characterized by the presence of well-stabilized secondary structural elements. (Brutscher et al. (1997); Cordier and Grzesiek (2004)) The relatively low λ_{noe} values observed for the non-native C-terminal helix, which has been shown to have

no long-lived long-range contacts with the other secondary structural elements, indicates that these λ_{noe} measurements are mainly sensitive to short-range local order. Finally, the U-state data show some type of "bell-shaped" curve reflecting the tumbling properties of polypeptide chains in solution. This feature has also been seen from other NMR data such as relaxation rates and residual dipolar couplings (Louhivuori et al. (2003)). Except for the terminal residues, all λ_{noe} values are between 0.4 and 0.7, indicative of the absence of any persistent local structure. Nevertheless, the variations observed along the polypeptide chain may be indicative of differential residual structure. Especially the segment comprising residues 36 to 44 displays slightly reduced λ_{noe} values. This segment comprises the two adjacent proline residues P37 and P38. It has been shown that proline residues are restricting the conformational freedom of the polypeptide chain, thus inducing some local order in the vicinity of these residues (Bernado et al. (2005))

Figure 8.8 also shows the results of 2D HET$_{ex}$-SOFAST-HMQC experiments performed on the same ubiquitin samples. These values do not provide useful information for the unfolded state due to the low pH and the slow intrinsic exchange rates. Very similar profiles are obtained by the λ_{ex} and λ_{noe} measurements, reflecting the close relationship between structural compactness and solvent accessibility. Because of the low pH of 2.0, hydrogen exchange for amides in the A-state and U-state samples is much too slow to yield significant variations in the λ_{ex} values.

8.5 Conclusions

We have presented HET-SOFAST NMR, a simple and robust method for the characterization of structural compactness in polypeptide chains. Application of the method to a selection of proteins comprising partially or fully unfolded as well as folded proteins shows that the method is able to distinguish reliably between these types of polypeptides in 1D experiments without the need of isotope labeling. The data obtained on α-lactalbumin in the apo- and Ca^{2+} -bound holo form demonstrate that the presented method is able to detect very subtle changes in the stability of the fold. If a ^{15}N-labeled protein sample is available, the 2D version of the experiments yields site resolved information providing insight into the structural heterogeneity of the polypeptide chain. If no resonance assignment is available this information is useful for the design and selection of optimized constructs of proteins being subject to structure elucidation by either NMR or X-ray crystallography. 2D HET-SOFAST HMQC may also be used for a fast confirmation of the presence and localization of secondary structural elements and loop regions in proteins with known ^{1}H-^{15}N resonance assignment.

The high sensitivity of the experiment allows data recording in a short experimental time even for low concentrated protein samples (μM). For higher sample concentrations (mM), the short acquisition times of 2D HET-SOFAST-HMQC may also allow real-time NMR measurements of changes in structural heterogeneity and solvent accessibility during protein folding or other kinetic processes occurring on a time scale of seconds. In addition, the availability of two complementary observables λ_{noe} and λ_{ex} prevents from an erroneous interpretation of the data, as changes in sample conditions and molecular structure have a distinguishable effect on these two parameters. Therefore we believe that HET-SOFAST NMR will prove a valuable new tool for the ongoing efforts in large-scale structure determination of proteins and protein complexes.

9 Protein folding and unfolding studied at atomic resolution by fast 2D NMR spectroscopy

9.1 Introduction

A detailed description of the structural changes occurring during the folding and unfolding of proteins still remains a challenging objective in biophysics. The understanding of the fundamental mechanisms of the folding process will also shed light on the factors leading to protein misfolding responsible for neurodegenerative diseases such as Alzheimer's and Parkinson's disease (Dobson (2004)). The ideal method to study protein folding/unfolding provides structural information at atomic resolution, and is sensitive to changes occurring on time scales ranging from microseconds to minutes, a task that no single technique is able to fulfill.

Nuclear Magnetic Resonance (NMR) spectroscopy is especially well adapted to obtain detailed atomistic information about the mechanisms, kinetics, and energetics of the folding/unfolding process for virtually every nuclear site in the protein. NMR methods are sensitive to molecular dynamics occurring over a wide range of time scales (Fig. 9.1a). While steady-state NMR methods are well suited to characterize equilibrium molecular dynamics occurring on a sub-second time scale (Mittermaier and Kay (2006); Palmer (2004)), unidirectional processes are best studied by real-time NMR (Van Nuland et al. (1998); Zeeb and Balbach (2004)) where a series of NMR spectra is recorded during the reaction (Fig. 9.1b).

Two difficulties have hampered so far the widespread application of real-time NMR to the study of protein folding. The first problem is the low intrinsic sensitivity of NMR at ambient temperature, a consequence of the small transition energies involved, compared to other spectroscopic techniques. Recent progress in NMR technology (magnetic field strength, cryogenic probes, electronics,...), how-

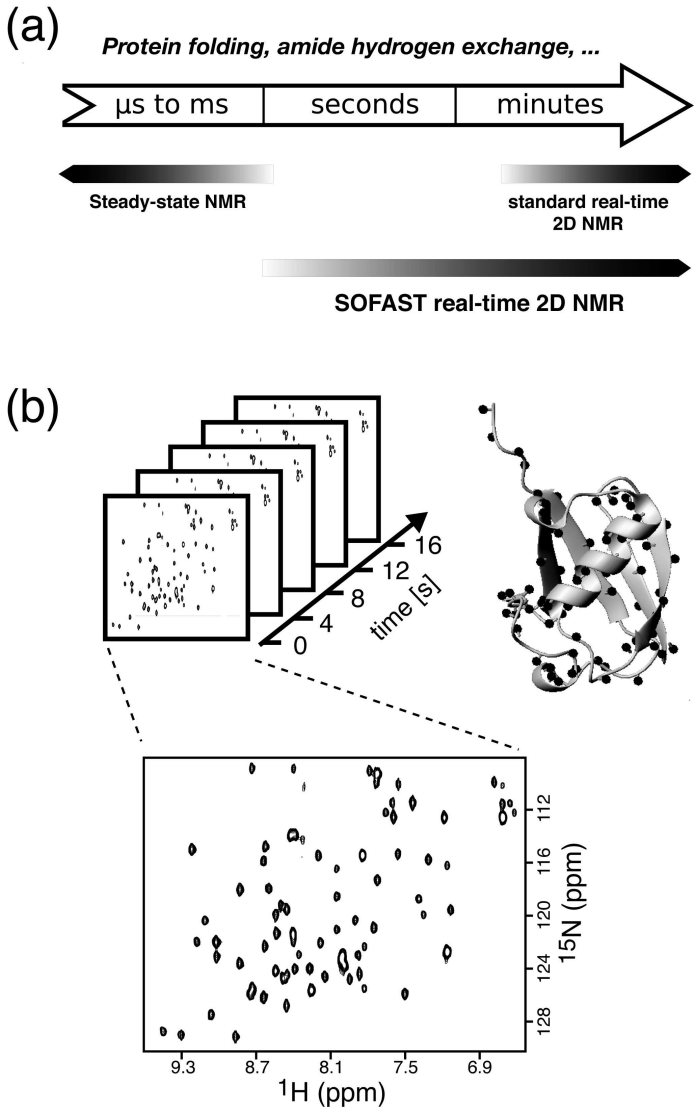


Figure 9.1 (a) Time scales of biophysical processes such as protein folding or amide hydrogen exchange, and NMR methods available to study the kinetics of these processes at atomic resolution. (b) Principle of SOFAST real-time 2D NMR. A series of 2D FTA-SOFAST- HMQC spectra is recorded after initiating a kinetic change in the protein state. Each cross peak reports on the local structure and dynamics at the site of a single amide group. The bottom spectrum has been recorded in an experimental time of 4 s on a 0.2 mM ^{15}N-labeled sample of ubiquitin at a magnetic field strength of 18.8 T using a cryogenically cooled probe.

ever, has greatly enhanced the sensitivity of NMR experiments, allowing acquisition of 1D spectra of proteins at sub-millimolar concentration at $\sim$ 1 to 10 s^{-1} rates. The second limitation concerns the time requirements to record higher-dimensional ($\geq$ 2D) NMR spectra indispensable to resolve the large number of nuclear resonances in a protein. Whereas a 1D NMR spectrum can be obtained in less than a second within a single scan by physical detection of the NMR time-domain signal induced in a receiver coil, the recording of a 2D NMR data set requires numerous repetitions (typically $\sim$ 100) of the basic pulse sequence in order to sample the evolution of the nuclear spins in the additional time dimension (Ernst et al. (1987)). This has limited in the past the application of real-time 2D NMR methods to relatively slow kinetic processes with characteristic time constants of minutes to hours. A major challenge for real-time NMR is therefore to achieve reduced acquisition times allowing to extend real-time NMR methods to the missing NMR time window of seconds to minutes (Fig. 9.1a). This will make the technique applicable to the study of a larger variety of molecular kinetics. Here we introduce SOFAST real-time 2D NMR, a method that allows site-resolved studies of kinetic processes in proteins with a time resolution down to a few seconds. We demonstrate the potential of this method for the study of protein folding and unfolding reactions.

9.2 Results and Discussion

9.2.1 SOFAST real-time 2D NMR spectroscopy.

To overcome the inherent time limitation of 2D NMR spectroscopy, two conceptually different approaches can be envisaged. First, the number of scans required to sample the multidimensional time space may be reduced using spectral aliasing, non-linear data sampling techniques, spatial frequency encoding, or Hadamard-type frequency-space spectroscopy. Second, accelerating the recovery of spin polarization between scans by means of optimized pulse sequences or the addition of relaxation agents to the solution, allows higher repetition rates of the pulse sequence. Here, we propose a fluid-turbulence-adapted (FTA) version of ^{1}H-^{15}N 2D SOFAST-HMQC to follow a kinetic process for individual amide sites in a protein (figure 9.2). Longitudinal relaxation optimization in FTA-SOFAST HMQC allows the use of scan times of less than 100 ms without a significant reduction in sensitivity (signal to noise per unit time), resulting in experimental times on the order of 5 seconds per 2D ^{1}H-^{15}N correlation spectrum, which shows one correlation peak for each non-proline residue of the protein. The advantage of the SOFAST approach with respect to other fast 2D NMR techniques, such as Hadamard or spatial frequency encoding, is the higher sensitivity provided by this technique for

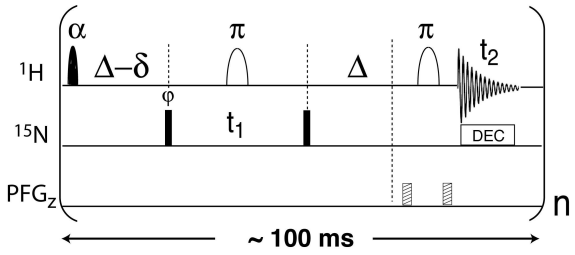


Figure 9.2 FTA-SOFAST-HMQC experiment used for SOFAST real-time 2D NMR measurements. The small angle ($\alpha < 90°$) ^{1}H pulse (filled symbol) is applied with a PC9 shape (Kupce and Freeman (1994)), while for the ^{1}H 180° pulses (open symbol) an iSNOB-5 shape (Kupce et al. (1995)) is used. The α angle is typically set between 40° and 60° depending on the scan time chosen. The transfer delays Δ are set to $1/(2J_{HN})$, and the small delay δ takes into account spin evolution during the PC9 pulse, and is adjusted to half of the PC9 pulse length. For ^{15}N chemical shift editing in t_1, the basic sequence is repeated n-times using short inter-scan delays (between 1 and 50 ms) resulting in scan times of $\sim$ 100 ms. Heteronuclear decoupling is performed using WURST-40 (Kupce and Freeman (1995)). The pulsed field gradients (PFGz) are applied for 100 μs with a strength of typically 30 G/cm.

the short acquisition times of a few seconds. High quality 2D spectra are obtained by the FTA-SOFAST-HMQC experiment on a modern high-field NMR spectrometer for protein samples in the sub-millimolar concentration range, as illustrated in Fig. 9.1b for a 0.2 mM sample of ubiquitin recorded in 4 s acquisition time. If faster sampling of the kinetic time domain is desired, this minimal experimental time required for FTA-SOFAST-HMQC can be further reduced to about 1 second using computer-optimized spectral aliasing in the ^{15}N dimension (Lescop et al. (2007), page 219).

In order to fully benefit from the increased time resolution provided by the FTA-SOFAST-HMQC experiment for real-time kinetic studies, a device is required that allows initiation of the kinetic event inside the magnet. The triggering of kinetic reactions can be achieved either by a sudden change of the protein state itself, e.g. by a photo-induced excitation or cleavage of chemical bonds (Rubinstenn et al. (1998); Wenter et al. (2005)), or by a change of environment, such as the solvent composition, temperature, or pH (Buevich et al. (2000); Zeeb and Balbach (2004)). Here we opted for an initiation of the reaction by an abrupt change in the solvent conditions achieved by rapid mixing of two solutions inside the NMR magnet (Mok et al. (2003); Van Nuland et al. (1998)). The fast mixing device used for this study (figure 9.7a on page 207) allows complete mixing in less than 300 ms as demonstrated by dye injection experiments into water solutions (figure 9.7b). A drawback of the fast injection is the presence of turbulences of the liquid over times significantly longer than the mixing time. Many modern NMR experiments

make use of pairs of pulsed field gradients for spectral artifact suppression. In the presence of bulk liquid motion incomplete refocusing of spin coherence due to translational diffusion during the time delay between the two pulsed field gradients yields uniformly reduced NMR signal intensities. The FTA-SOFAST-HMQC experiment has been optimized for minimal signal loss in the presence of bulk liquid motion, yielding dead times after the injection of less than two seconds (figure 9.8).

We have chosen two conceptually different applications in order to demonstrate the potential of SOFAST real-time 2D NMR to provide accurate measures of kinetic rate constants for a large number of amide sites: The first example concerns the folding of α-lactalbumin from a molten-globule to the native state. In the second application, the unfolding kinetics experienced by individual residues of ubiquitin under native equilibrium conditions are investigated.

9.2.2 Conformational transition kinetics of α-lactalbumin from a molten-globule to the native state studied in real time.

The protein α-lactalbumin serves as a model system to investigate the structural transition from a partially folded to the native state using SOFAST real-time 2D NMR. This protein has been extensively used as a model system for protein folding studies. The structure of native α-lactalbumin (14 kDa) comprises two domains, one containing four α-helices and a short 3_{10} helix, and the other one containing a three-stranded β-sheet and another 3_{10} helix. Here, we focused on the Ca-free (apo) form of α-lactalbumin. Under destabilizing conditions such as low pH, addition of co-solvents, high temperature, or combinations thereof, apo α-lactalbumin exists in a molten-globular (MG) state. The MG state is characterized by the absence of long-lived tertiary structure, but it still contains a high degree of secondary structure (Mok et al. (2005); Troullier et al. (2000)), with a radius of gyration only $\sim$10% larger than the native state (Balbach (2000); Kataoka et al. (1997)). The MG state of α-lactalbumin has recently attracted considerable interest when it was realized that it may act as an important component causing apoptosis of tumor cells (Gustafsson et al. (2004)).

Pioneering work on the use of time-resolved NMR methods for monitoring protein refolding was performed on α-lactalbumin (Balbach (2000); Balbach et al. (1996, 1995); Forge et al. (1999); Van Nuland et al. (1998)). 1D ^{1}H NMR spectra recorded during the refolding reaction of apo α-lactalbumin provided evidence that the MG state, accumulated during the early stages of the folding reaction, shows similar spectral characteristics (poorly dispersed and broad resonances),

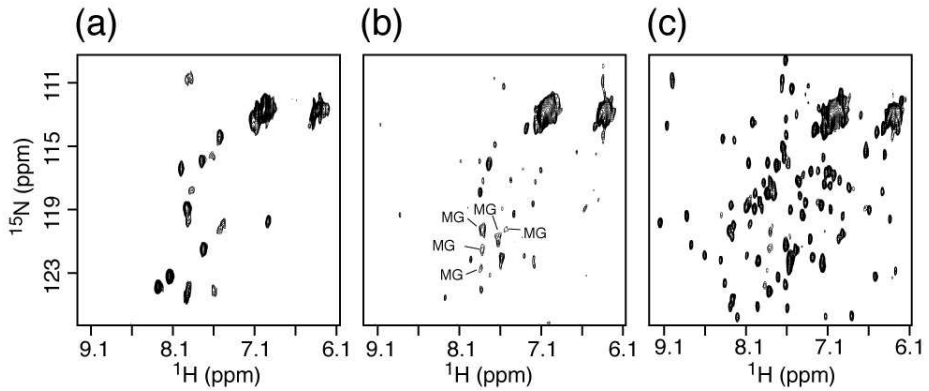


Figure 9.3 FTA-SOFAST-HMQC spectra of bovine α-lactalbumin at pH 2.0 (left spectrum), immediately after a sudden pH jump to pH 8.0 that triggers folding (center spectrum) and 120 s after injection (right spectrum). Each spectrum shows the sum of two acquisitions of 10.9 s duration. Peaks corresponding to the MG state that disappear during folding are annotated.

and therefore similar dynamical properties of the conformational ensemble, as those observed for the MG state stabilized at acidic pH (Balbach et al. (1995)). Additional diffusion-edited and nOe-transfer NMR measurements allowed monitoring the compactness of the protein and the establishments of native tertiary contacts, respectively (Balbach (2000); Forge et al. (1999)). The folding kinetics of apo α-lactalbumin have also been studied previously at a residue level by line shape analysis of individual cross peaks detected in a single 2D NMR spectrum recorded during the refolding event (Balbach et al. (1996)). This technique, however, is limited to relatively slow kinetics (several minutes), and is prone to large experimental uncertainties. In the study of Balbach et al. (1996), folding rates for a total of 25 backbone amides could be quantified with an estimated experimental error of $\sim$25%. Because of the small number and the high uncertainties of the measured rate constants, these data did not allow to definitely exclude the presence of differential folding kinetics along the polypeptide chain.

In order to illustrate the improvements provided by SOFAST real-time 2D NMR for the measurement of residue-specific folding rates, we monitored the same refolding reaction as Balbach et al. from the MG state, formed at pH 2, to the native state. Refolding was initiated by a sudden pH jump (from 2 to 8), and the reaction was monitored by a series of FTA-SOFAST-HMQC spectra recorded at a 0.1 s^{-1} rate. Spectra acquired before injection, immediately after, and 2 min after injection are shown in Fig. 9.3. The central spectrum provides a snapshot of the transient MG state under native conditions. The broad and weak signals in this spectrum are indicative of large-scale dynamics occurring on a micro- to millisecond timescale. These data are in agreement with this transient MG-state being a compact, highly

dynamic ensemble of conformational states. All NMR signals observed during the folding process can be assigned either to the MG or the native state. No additional peaks indicative of a significantly populated folding intermediate were detected in agreement with previous findings (Balbach (2000); Balbach et al. (1996); Forge et al. (1999)). The refolding kinetics could be quantified for a total of 92 out of 121 backbone amide sites in the protein from intensity measurements of well- resolved cross peaks in the ^{1}H-^{15}N correlation spectra (Fig. 9.4). In addition, for the first time the intensity decay of 5 cross peaks characteristic for the MG state could be quantified. Under the experimental conditions chosen (15°C, pH 8, Ca^{2+} -free), the NMR intensity decay and buildup curves can be fitted to mono-exponential functions. Fitting the data to more complex e. g. bi-exponential or stretched exponential functions does not yield statistically significant improvements. The folding time constants measured for individual amide sites of the native protein and the MG state are identical (τ_f = 109 $\pm$ 5 s) within the experimental uncertainty. The obtained folding kinetics are consistent with results from fluorescence measurements (figure 9.5), yielding a characteristic time constant of 110 $\pm$ 10 s for the establishment of native tertiary structure. The good agreement with the fluorescence measurements demonstrates the accuracy obtained by SOFAST real-time 2D NMR experiments.

The MG state of apo α-lactalbumin explores a large conformational space, possibly including the presence of native and non-native secondary structures (Troullier et al. (2000)). Therefore the transition from this MG state to the native state requires a large decrease in conformational entropy, known as entropy bottleneck effect (Bushmarina et al. (2006); Wolynes et al. (1995)). This is the main reason for the significantly slower folding kinetics observed for the apo form compared to the holo form of α-lactalbumin (Bushmarina et al. (2006)). The high precision of rate constants (error of $\sim$5-10%) obtained in our study for a large number of individual amide sites allows to draw some conclusions on the energy landscape and folding pathways. The finding of equal kinetic rates for the buildup of the native state, and the disappearance of the MG state indicates a transition between only two states, the MG ensemble representing a large number of conformational sub-states interconverting on the micro- to millisecond time scale, and the native state. In other words, the establishment of the native tertiary structure is not accompanied by a change in the structural and dynamic properties of the MG ensemble as folding proceeds. These observations are in agreement with the assumption of a smooth energy landscape where the folding rate is controlled by a single transition state ensemble originating from the conformational entropy bottleneck effect. The existence of a unique transition state is in agreement with conclusions drawn from other biophysical folding experiments (Bushmarina et al. (2006)).

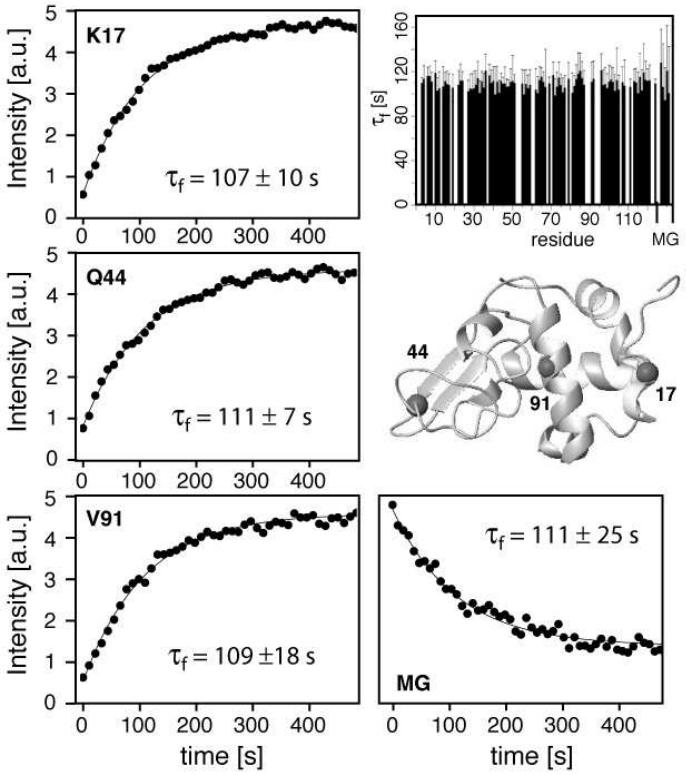


Figure 9.4 Refolding kinetics of bovine α-lactalbumin from the MG-state to the native state. The measured peak intensities are plotted as a function of the folding time. Shown are three residues situated in loop (K17), β-sheet (Q44) and α-helical regions (V91) that are indicated on the structure (PDB entry 1F6R). In addition, the signal decay observed for a peak assigned to the MG state is shown. Solid lines represent best fits to a three-parameter exponential function. A histogram shows the measured folding time constants for 92 residues in the native state, as well as the 5 rates measured for the disappearance of the MG state.

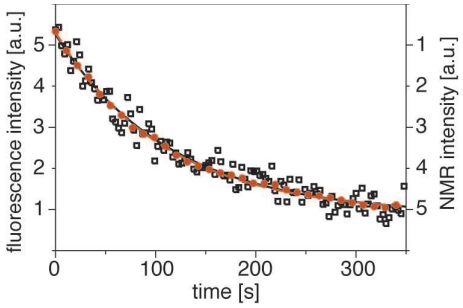


Figure 9.5 The refolding kinetics of bovine α-lactalbumin studied by fluorescence (squares) is compared to the intensity buildup of the amide ^{1}H-^{15}N signal of residue 51 in SOFAST real-time 2D NMR (red circles). For reasons of better comparability the buildup of NMR intensity is plotted from top to bottom. For experimental details see the Methods section (page 207).

9.2.3 Ubiquitin unfolding kinetics under equilibrium conditions revealed by EX1 amide hydrogen exchange measurements.

To further illustrate the potential of SOFAST real-time 2D NMR to detect differential kinetic behavior along the polypeptide chain, we have investigated the unfolding kinetics of human ubiquitin under non-denaturing equilibrium conditions by amide hydrogen-deuterium (H/D) exchange methods. H/D exchange is a powerful tool for the study of protein structure, folding, unfolding, and binding (Ferraro et al. (2004); Krishna et al. (2004)). The exchange between amide and solvent hydrogens requires that the amide group is in an exchange-competent, solvent-accessible conformation. This can be achieved by local, sub-global or global protein unfolding fluctuations (opening reactions) that transiently break the exchange-protecting structures, e.g. hydrogen bonds. The measurement of amide H/D exchange rates thus provides residue-specific information on the otherwise invisible manifold of partially or globally unfolded excited conformational states populated at a very low level. Despite their low population under native conditions, such high energy states may be crucial for protein function (Eisenmesser et al. (2002); Frauenfelder et al. (1991)), or they may represent intermediate states on the protein folding pathway (Neudecker et al. (2006)). Especially interesting are H/D exchange measurements in the so-called EX1 regime, where the measured exchange rate constants directly reflect the kinetics of the unfolding reaction at individual amide sites (Ferraro et al. (2004)). EX1 conditions are generally reached at high pH. However, for most proteins, H/D exchange at high pH is too fast to measure by conventional 2D real-time NMR. Therefore previous EX1 studies had to focus on a small number of well-protected amide sites. It has been shown that the EX1 exchange rates measured for a few slowly exchanging amide sites in ubiquitin correspond to the rate of global unfolding observed with other biophysical methods (Sivaraman et al. (2001)). The SOFAST real-time 2D NMR method, presented here, allows to extend H/D exchange measurements under EX1 conditions to a significantly larger number of amide sites, yielding a more comprehensive picture of the potentially heterogeneous unfolding processes along the polypeptide chain. Representative exchange curves measured for the 76-residue protein ubiquitin at pH 11.95 and 25°C are shown in Fig. 9.6. For all amides, even in the most protected parts of the protein, exchange time constants $\tau_{ex} < 200$ s were observed. The high quality of the data allowed accurate quantification of exchange kinetics down to $\tau_{ex} \sim 5$ s, with a detection limit of $\tau_{ex} \sim 1$ s. This proved to be sufficient to measure exchange rates for amides in secondary structural elements, whereas H/D exchange in loop regions was generally completed during the dead time of the experiment. In the upper drawing of Fig. 9.6, the measured exchange rates are color-coded on the ribbon structure of ubiquitin. The slowest H/D exchange

is observed for residues in strands I and II of the β-sheet (residues 3-5, 15), and the central part of the α-helix (residue 27), while amide hydrogens located in other parts of the molecule, comprising strands III to V, exchange much faster. Interestingly, a gradual increase in the kinetics of opening fluctuations across the β-sheet is observed, indicating non-cooperative unfolding events in this β-sheet, reminiscent of a hydrophobic zipper folding mechanism (Dill et al. (1993)).

Ubiquitin has been used over the past 20 years as a model system for the study of protein stability and folding using a variety of biophysical techniques, denaturation methods, protein engineering studies, and computational approaches (Jackson (2006)). Therefore, it is interesting to compare the results presented here with previous findings on the thermodynamic stability, folding and unfolding kinetics of ubiquitin under different experimental conditions. H/D exchange rates measured under EX2 conditions (Bougault et al. (2004); Pan and Briggs (1992)) provide a measure of the thermodynamic equilibrium constant $K = k_{open}/k_{close}$, rather than a kinetic rate constant as obtained under EX1 conditions. Interestingly, the residue-specific equilibrium constants also show differential stability among the secondary structural elements of ubiquitin, with notably the first two β-strands and the α-helix being the most stable parts. The kinetic information obtained from H/D exchange measurements under EX1 conditions complements these former results, indicating that the differential thermodynamic stability observed for ubiquitin is mainly determined by the unfolding kinetics (opening rates). Differential non-cooperative unfolding has also been observed by NMR studies of ubiquitin under denaturing conditions. The populations of non-native conformational states can be enhanced by changing the temperature, pressure, or pH, or by adding alcohol co-solvents. A cold-denaturation study by Wand and co-workers shows for reverse-micelle-encapsulated ubiquitin that the mixed β-sheet is progressively destabilized from the C-terminal side in the temperature range from -20° to -30°C (Babu et al. (2004); Pometun et al. (2006)). In a NMR study by Cordier and Grzesiek (Cordier and Grzesiek (2002)), the change in hydrogen bond strength with increasing temperature was measured. The N-terminus of β-strand V was found to be the least thermally stable, whereas strands I and II remain stable even at elevated temperatures. Spectral changes observed upon high-pressure denaturation at ambient temperature indicate the presence of a partially unfolded intermediate in which the C-terminal part (residues 70-76) as well as residues 33-42 and residue 8 are denatured (Kitahara and Akasaka (2003)). The higher stability of the secondary structural elements in the N-terminal part of ubiquitin (β_I-β_{II} strand and α-helix) is also evidenced by the observation of a partially structured state (A state) at low pH in a 60%/40% methanol/water mixture (Wilkinson and Mayer (1986)). In the A-state, the N-terminal part of ubiquitin retains native-like secondary struc-

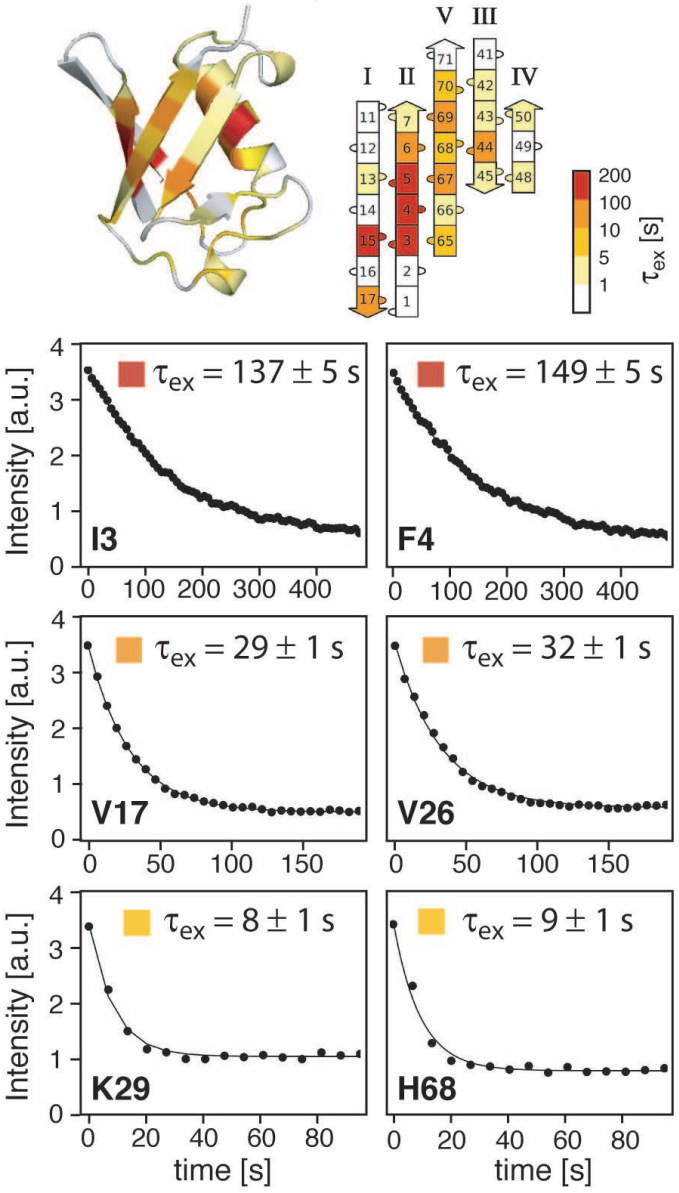


Figure 9.6 Hydrogen/deuterium exchange data obtained for human ubiquitin at pH 11.95 using SOFAST real-time 2D NMR. The measured exchange rates are color coded on the ubiquitin structure in the top panel. Examples of exchange curves corresponding to different exchange regimes are shown in the lower panels together with the fitted exchange time constants. For residues color-coded in white, no signal decay was observed as the cross peak intensity has decreased to its plateau value during the dead time of the experiment. In the schematic representation of ubiquitin's β-sheet, the orientation of the N-H bonds is indicated by half-circles.

ture, whereas the C-terminal part undergoes a conformational transition form the native state to an ensemble of conformational states with high helical propensity (Brutscher et al. (1997)). Finally, a recent time-resolved infrared (IR) study reveals a complex thermal unfolding behavior of ubiquitin spanning a wide range of time scales (Chung et al. (2005)), indicating a gradual, rather than a cooperative, unfolding process. All experimental data from both equilibrium and transient unfolding studies support the picture of non-cooperative unfolding events taking place in the native environment. The C-terminal part, comprising β-strands III to V, is significantly less stable, with unfolding rates that are up to 2 orders of magnitude higher than those observed for β-strands I and II, and the α-helix in the N-terminal part. The observed conformational dynamics under native conditions, and the presence of partially unfolded conformational states may be of importance for the biological function of ubiquitin, as it has been shown that incorporation of a disulfide bridge between residues 4 and 66, stabilizing the connection between strands I and V of the β-sheet, leads to a 70-80% decrease of activity in signaling proteolysis (Ecker et al. (1989)).

SOFAST real-time 2D NMR closes the gap (seconds to minutes) on the kinetic time scale (see fig.9.1). This makes possible H/D exchange measurements for a large number of amide sites at any desired pH by 2D NMR methods. The combined interpretation of H/D exchange rates measured under EX1 as well as the more common EX2 conditions presents an attractive new tool to access residue-specific kinetic folding and unfolding rates of proteins under native (physiological) conditions. In concert with other NMR data, providing the same atomic resolution, e. g. spin relaxation measurements or NMR spectra recorded under denaturing conditions, H/D exchange measurements will help to gain a deeper insight into the nature of partially unfolded conformational states present under native conditions, and the rates of interconversion between those excited states and the native state.

In summary, we have demonstrated that SOFAST real-time 2D NMR allows the accurate measurement of kinetic rate constants up to about 1 s^{-1} simultaneously for a large number of amide sites in a protein. The time frame ranging from a few seconds to roughly a minute used to be inaccessible to conventional 2D NMR methods. We have illustrated the potential of SOFAST real-time 2D NMR to provide accurate residue-specific kinetic information on protein folding and unfolding events either by direct real-time studies of folding/unfolding, or indirectly by real-time measurements of H/D exchange kinetics. The kinetic rates measured for the refolding of α-lactalbumin are uniform throughout the protein indicating a smooth energy landscape for the transition from the molten globule to the native

state. On the contrary, the kinetic data collected for the protein ubiquitin indicate a highly non-cooperative unfolding behavior under native conditions in agreement with previous results from NMR and infrared spectroscopy. The methodology presented here can be easily applied to protein systems with molecular weights ranging between ∼5 and ∼20 kDa where deuteration is not a prerequisite for obtaining well-resolved H- N correlation spectra. Here we have focused on the application of SOFAST real-time 2D NMR to the study of protein folding and unfolding, but the method also offers new opportunities for residue-specific kinetic measurements of other unidirectional events occurring on a time scale of seconds such as binding, enzymatic reactions, and chemical exchange. Combined with recent advances towards NMR studies of proteins inside intact living cells (Reckel et al. (2005); Selenko et al. (2006)), SOFAST real-time 2D NMR may also become a powerful experimental tool for in situ time- and site-resolved observation of kinetic events such as protein folding and posttranslational modifications in a cellular environment.

9.3 Methods

9.3.1 A fast injection device for real-time NMR studies: Principle, problems and solutions

A schematic illustration of the fast-injection device that allows rapid mixing of two solutions inside the magnet is shown in figure 9.7(a).

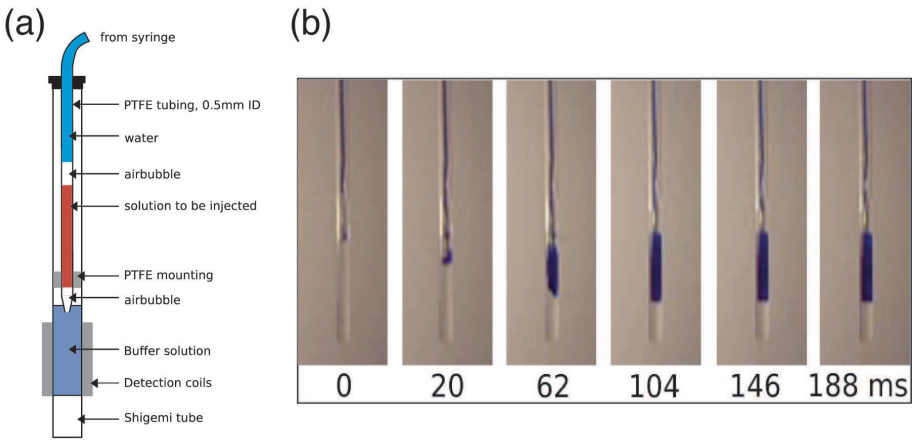


Figure 9.7 (a) Sketch of the fast injection device used for SOFAST real-time 2D NMR applications. (b) Photographs of the NMR sample tube taken after fast injection of 50 μL blue dye into 350 μL of water using a rapid injection device. Homogeneous mixing is completed within approximately 100 milliseconds.

When building this device it was found important to have strong turbulences, in order to achieve a homogeneous mixing. This could be obtained paying attention to a number of factors: (i) The velocity of the injection should be high. To achieve this, the bottom of the injection tube has a small diameter: Whereas the main part of the Teflon tube (which is connected to the syringe) has an inner diameter of 0.5 mm, its end diameter is smaller. In practice this was done by pulling the tube under heat, and then cutting the tube at the constricted part, where the diameter is smallest. One has to make sure the the solution is injected in the axis of the device (straight, see ii). (ii) The solution has to hit a plane surface in the NMR-tube, and should not hit the sidewalls, otherwise it just flows down on one side of the tube and is slowed down. To this end a Shigemi tube (flat bottom) was used, and the injection device was built so that the solution is injected straight to the bottom. The injection device was kept in the middle of the NMR tube using Teflon mountings (a Teflon tube with an outer diameter that closely matches the inner diameter of the NMR tube was fixed around the injection device). The use of a Shigemi tube has another advantage: The volume can be reduced to ca. 350 to 400 μL.

Figure 9.7(b) illustrates that with such an injection device homogeneous mixing is achieved in less than 100ms.

However, the turbulences of the solution, that are necessary for the fast homogeneous mixing, are found confining when using pairs of magnetic field gradients.

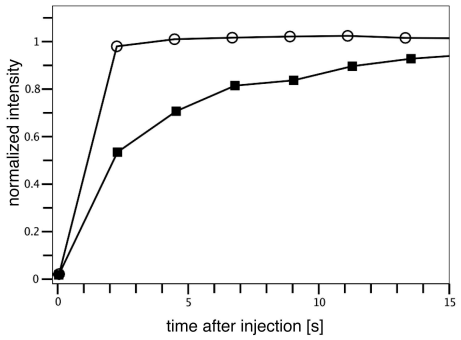


Figure 9.8 Fast injection and magnetic field gradients. See text for details.

To illustrate this, figure 9.8 shows the Gly-H^{α} signal intensity measured after rapid injection of glycine into D_2O for two different gradient echo sequences: 50 μL of a 2M glycine solution (in D_2O) were injected in 350 μL D_2O and the signal intensity of glycine's H^{α}-peak was followed as a function of time using a 1D gradient spin-echo pulse sequence (small-flip-angle pulse - PFG_z - $\tau/2$ - 180° pulse - $\tau/2$ - PFG_z - acquisition). This sequence is sensitive to translational diffusion effects, and

thus to the fluid turbulences created by the fast injection. It mimics the situation found in more complex NMR experiments such as SOFAST-HMQC. Results are shown for dephasing times of τ = 20 ms (filled squares), corresponding to the situation in standard SOFAST-HMQC, and τ= 1 ms (open circles), corresponding to the situation in fluid-turbulence-adapted FTA-SOFAST-HMQC. Stable conditions are reached in less than $\sim$ 2 s using the short dephasing time, while $\sim$ 15 s elapse until equilibrium is reached when using the longer delay, emphasizing the interest of FTA-SOFAST-HMQC when using a rapid injection device. This behavior can be expected to depend somewhat on the mixing device used and on the speed of the injection; nevertheless, the trend has been found reproducible when repeating this test.

9.3.2 SOFAST real- time 2D NMR

All NMR experiments were performed on a Varian INOVA spectrometer operating at 800 MHz ^{1}H frequency equipped with a cryogenic probe. To keep the dead time as short as possible the acquisition of the first spectrum was started after the dead time of the mixing required to stabilize the fluid turbulence, which was about two seconds (see above, figure 9.8). In order to achieve high time resolution while retaining good spectral quality without the need of advanced spectral processing we used alternate phase cycling in subsequent spectra as follows. Each experiment is performed with only one scan per increment in t_1, ensuring highest repetition rates of experiments, and the sign of the ^{15}N excitation pulse phase and receiver phase in FTA-SOFAST-HMQC is alternated between subsequent experiments. Whereas the sign of the ^{15}N encoded signal of the protein is retained in the two experiments, artefacts have opposite sign. Adding 2 succeeding spectra (1 and 2, 2 and 3, 3 and 4,...) along the kinetic time dimension removes spectral artifacts such as axial peaks or t_1 noise at the water frequency and improves the base line. Figure 9.9a demonstrates this approach for a H/D exchange experiment on ubiquitin. This allows a more accurate and precise measurement of spectral parameters, such as peak positions and intensities.

How does this procedure affect the kinetic information that can be extracted? Figure 9.9b investigates this issue for the case of a monoexponential function. The black solid line represents the exponential change in peak intensity that is sampled at a rate determined by the acquisition time of one spectrum, as indicated by black circles. Co-adding two successive spectra corresponds to replacing two data points by their linear average and assigning this summed value to the central time point in between the experimentally measured data points. These data points (which for reasons of clarity were scaled in the figure) are indicated as green circles and the linear approximation connecting the experimental data points is indicated by

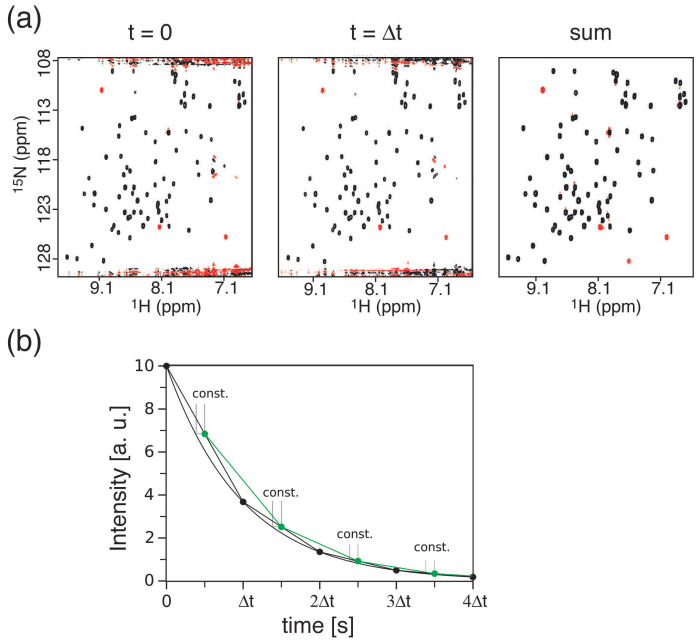


Figure 9.9 Alternate phase cycling in real-time 2D NMR. (a) Two spectra, acquired at times t =0 (left panel) and $t = \Delta t$ (central panel), where Δt is the duration of one 2D experiment, are acquired with one scan per t_1 increment. The phases of the first ^{15}N pulse and the receiver are alternated by 180° between the two experiments. The right panel shows the sum of these two spectra. (b) Analysis of the effect of this co-addition procedure on the extracted exponential decay rates.

black straight lines. The mono-exponential rate determined by these co-added data points corresponds exactly to the actual rate, which is explained by the fact that these data points are shifted by a constant offset independent of time. The only prize to pay is a longer dead time, as can be appreciated from the position of the first data point (green) in the co-added data set.

9.3.3 Real-time folding of bovine α-lactalbumin

Nitrogen-15 labeled bovine α-lactalbumin (M90V) was expressed and purified as described previously (Grobler et al. (1994)). The MG-state was prepared by dissolving 8.8 mg protein in 350 L H_2O/D_2O (10:1) and adjusting the pH to 2.0 with 1M HCl solution. Fifty μL of a refolding buffer solution (H_2O/D_2O: 10:1), containing 800 mM Tris and 80 mM EDTA, pH 9.1, were loaded in the injection device, and refolding was initiated inside the spectrometer by rapid injection into the protein solution. The folding reaction was followed by a series of FTA-SOFAST-HMQC experiments of 10.9 s duration using a scan time of 130 ms and 40 complex data points in t_1 (80 scans per 2D spectrum). The final pH measured after the folding

reaction was 8.0, the final protein concentration was 1.55 mM. NMR assignments were taken from the work by Forge et al. (1999). The reproducibility of the real-time folding data has been evaluated by repeating the experiment twice. The difference in the rate constants obtained from the two measurement series is smaller than the experimental error estimated from Monte Carlo simulations of a single data set.

The protocol used for the fluorescence measurements (protein concentration, pH-jump, solvent...) is identical to that used for the time-resolved NMR experiments, except for the initiation of the folding reaction. For the fluorescence measurements manual mixing yielded a dead time of about 1 minute. The excitation wavelength is 295 nm (bandwidth: 1 nm) and the fluorescence change is monitored at 340 nm (bandwidth: 3 nm). Non-optimal sample volume, and the resulting small detection window, as well as auto-quenching effects because of the high sample concentration explain the relatively low signal-to-noise ratio obtained by the fluorescence experiment. A high protein concentration was chosen to mimic the NMR conditions, resulting in a small volume for the available amount of 2 mg ^{15}N-labeled protein. The same recombinant ^{15}N-labeled α-lactalbumin has been used for NMR and fluorescence experiments because the recombinant protein (expressed in E. Coli) is less stable than the natural protein, and this has an effect on the refolding kinetics (Ishikawa et al. (1998)). The experimental data were fitted to mono-exponential decay.

9.3.4 Amide hydrogen exchange measurements in ubiquitin

For the SOFAST real-time 2D NMR amide hydrogen exchange measurements, 2.5 mg of human ubiquitin were dissolved in 50 μL of H_2O buffer containing 50 mM glycine and loaded into the injection device. H/D exchange was initiated by injection of the protein solution into 350 μL D_2O buffer, resulting in a final protein concentration of 0.7 mM. The resulting pH, corrected for the electrode isotope effect was 11.95. The decay of peak intensities due to exchange was followed by recording a series of FTA-SOFAST-HMQC spectra of 6.7 s duration using a scan time of 112 ms, and recording 30 complex data points (60 scans). NMR assignments at pH 11.95 were obtained from published data at pH 6.6 (Cornilescu et al. (1998)), and a series of ^{1}H-^{15}N SOFAST-HMQC spectra recorded in the pH range 7 to 12.

9.3.5 Validating the correctness of extracted H/D exchange rates from SOFAST-HMQC experiments

SOFAST NMR makes use of longitudinal amide ^{1}H relaxation enhancement to increase the experimental sensitivity for high repetition rates of the experiment. Shorter spin-lattice relaxation times are achieved by the use of amide-proton selective radio-frequency pulses that leave the polarization of other ^{1}H spins, e. g. aliphatic and water ^{1}H, unperturbed. The energy put into the system is then efficiently dissipated within the dipolar-coupled ^{1}H spin network by spin diffusion (nOe) effects. In the course of H/D exchange, amide hydrogens are progressively exchanged for deuterons. This alters the ^{1}H spin coupling network in the protein from one 2D experiment to the other, which influences the relaxation properties of the amide ^{1}H spins. How does this change in the ^{1}H network change the peak heights, and thus the extracted H/D exchange rates ?

In a non-selective experiment the exchange of amide protons for deuterons should have negligible effects on the intensity of the remaining amide protons: The relaxation after non-selective excitation is governed by the evacuation of excitation energy from the whole spin system (see the discussion in section 2.3.2.2, e.g. figure 2.14). Removal of amide protons during H/D exchange means only a small *relative* change in the total number of ^{1}H spins present, and therefore only a negligible *relative* change of the excitation energy present after non-selective excitation. The relaxation of the remaining amide protons is therefore almost not changed by the exchange of some amide protons for deuterons. This is confirmed by simulations (data not shown).

In selective experiments however, the situation is different: The energy that has to be evacuated from the spin system is determined by the number of *amide* protons present in the protein, because only these are excited. The excitation energy present in the spin system at the beginning of H/D exchange (when all amide sites are still protonated) may therefore be much larger then at a later time point, when many amide sites are deuterated. The longitudinal relaxation of a given (slowly exchanging) amide ^{1}H spin may therefore significantly change during H/D exchange. The peak intensities measured at different times may therefore not correctly reflect the H/D exchange process of such a proton, but may be due to the changing relaxation properties, and the extracted H/D exchange rates in SOFAST HMQC experiments would be incorrect.

We have investigated the question of accuracy of exchange rates determined from SOFAST HMQC by experiment and computer simulation. Figure 9.10a shows a comparison of exchange rates obtained from a series of se-wfb HSQC spectra and SOFAST HMQC experiments. As the rates of the former should not (significantly)

be influenced by the changing ^{1}H network, and thus reflect the *true* kinetic rates, this comparison is a measure of the accuracy of the SOFAST HMQC experiments.

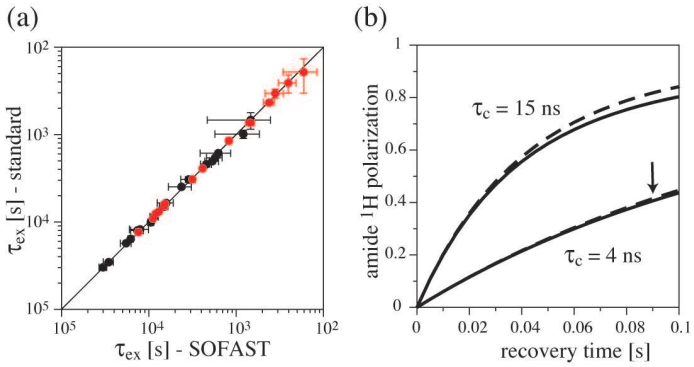


Figure 9.10 Evaluation of the accuracy of measured H/D exchange time constants using SOFAST real-time 2D NMR. (a) τ_{ex} values measured using the FTA-SOFAST-HMQC sequence of Fig. 9.2 are compared to results obtained using standard HSQC-based 2D NMR methods. H/D exchange was monitored during 24 hours for ubiquitin at 25°C at two different pH values : 7.3 (black circles), and 9.8 (red circles). HSQC (expt. time: 3 min 15 s, recycle delay: 0.9 s) and FTA-SOFAST-HMQC (expt. time: 25 s, recycle delay: 50 ms) spectra were recorded in an interleaved manner resulting in a 0.14 min sampling rate of the exchange process. H/D exchange was initiated by fast injection of 50 μL of a H_2O solution into 350 μL D_2O solution. Under these experimental conditions, the data recorded at pH 9.8 (red circles) represent the 15 most slowly exchanging amides, whereas the data measured at pH 7.3 allowed quantifying the exchange rates for less protected amide sites. (b) Computer simulation of the relaxation behavior of a given amide proton when all other amide sites are protonated (solid curves) and when all other amide sites are deuterated.

The good correlation of the data obtained with the two methods (correlation coefficient of R=0.997) shows that the changing ^{1}H network does not have significant effect on the H/D exchange rates from SOFAST-HMQC experiments.

The experimental results shown in (a) may be rationalized by computer simulations, shown in figure 9.10(b). The buildup of amide ^{1}H polarization has been calculated for ubiquitin by numerical integration of the Solomon equations, as described in section 2.3, assuming a magnetic field strength of 18.8 T (800 MHz ^{1}H frequency), and different rotational correlation times τ_c of the protein. Two extreme cases have been considered: (i) all backbone amide sites are protonated (solid curves), and (ii) only the observed amide site is protonated while all others are deuterated (dashed curves). The result of these simulations is that polarization recovery is slightly faster for the second case. This can be understood by the fact that fewer amide protons are excited in the molecule, and therefore less energy has to be evacuated from the spin system to restore the equilibrium polarization. Nevertheless, for the fast pulsing conditions used here (indicated by an arrow),

and the tumbling correlation time of ubiquitin at 25°C (τ_c=4 ns), the difference in the restored polarization is negligible ($< 1\%$). For larger proteins the difference increases (see curves obtained for τ_c= 15 ns) because of the more efficient spin diffusion. However, if higher repetition rates (shorter recovery times) are used for such large proteins, the difference becomes again negligible.

10 Appendix

10.1 Algorithm for simulation of longitudinal relaxation in a multispin system.

To get more insight into proton relaxation in proteins, a simulation program was written (in Python language), based on equation 2.38 to follow the time course of the polarization of every proton spin in a molecule. The program is based on the following principles:

- The coordinates of the proton spins and the internuclear distances are read from a PDB file (using the package BioPython, http://biopython.org)
- For the calculation of the relaxation rate constants ρ and σ the spectral density function is modeled with a simple model: overall tumbling is assumed to be isotropic with a correlation time τ_c. Internal dynamics is ignored except for methyl rotation where a rotation around the methyl axis is assumed. With these assumptions, interactions between non-methyl protons and interactions between a methyl proton and a non-methyl proton are modeled using equation 2.33 with the following expression for the spectral densities:

$$J(\omega) = \frac{\tau_c}{1 + \omega^2 \, \tau_c^2} \tag{10.1}$$

  Interactions within a methyl group were modeled with a generalized order parameter S^2, which describes the rotation of the methyl group around the symmetry axis, and a correlation time τ_i for this rotation (yielding an effective correlation time $\tau_e^{-1} = \tau_c^{-1} + \tau_i^{-1}$). The order parameter was set to 0.11, which reflects the covalent geometry of the methyl group (Woessner (1962)) and the correlation time constant for methyl rotation τ_i was set to 50 ps, which is on the order of experimentally found values (Urbauer et al. (1996)).

$$J(\omega) = \left(\frac{S^2 \tau_c}{1 + \omega^2 \, \tau_c^2} + \frac{(1 - S^2)\tau_e}{1 + \omega^2 \, \tau_e^2} \right) \tag{10.2}$$

- The time evolution of the spin system was calculated by straightforward multiplication of equation 2.38. The time derivative $\mathrm{d}\,(I_z - I_z^0)/\mathrm{d}t$ was calculated in intervals using a step-size of approximately 0.5 ms.
- Water exchange effects were accounted for as follows: Direct exchange of amide protons with water was ignored. This is justified by the fact that in compact secondary structures in proteins exchange is generally much slower than the time scale of milliseconds that is relevant for the longitudinal relaxation. The exchange of the solvent-exchangeable sidechain hydrogen atoms (Asn Hδ, Gln Hϵ, Arg Hϵ and Hζ, Cys Hγ, His Hϵ_2 and Hδ_1, Tyr Hζ, Ser Hγ and Thr Hγ_1) was assumed to be very fast. This is justified by experimental results (Wider (1998)). In the simulation their state of polarization was set equal to the polarization state of the water proton spins at each time point of the calculation. In the simulations of water flip-back experiments (i.e. where the water magnetization is aligned along +z at the beginning of the recycle delay) the polarization of these sidechain atoms was therefore held constant along +z. In simulations where water is excited a longitudinal relaxation time of 3 seconds was assumed for water and thus also for these fast-exchanging protons. Any alteration of the bulk water proton polarization by this exchange can safely be neglected, as the concentration of water is on the order of 50000 times higher than the concentration of protein. The water ^{1}H relaxation time of 3 seconds is based on experimental observations (Hiller et al. (2005b))
- The highly non-monoexponential nature of relaxation in multispin systems precludes the description of the recovery by a single T_1 value. In cases where "relaxation times" are reported these were determined as follows. If relaxation was monoexponential with a characteristic time constant T_1 then after 90° excitation the polarization would reach a level of $\approx 63\%$ (i.e. $1 - 1/e$) after a time T_1. Based on this and in order to reflect the relaxation process as one single number the reported effective T_1 times refer to the time when 63% of the longitudinal magnetization is restored after 90° excitation.
- For the relaxation simulations on ubiquitin, the very dynamic four C-terminal residues were removed before simulation. Because of their dynamics the model is found too simplistic to reproduce their relaxation behavior (see figure 2.21).

Despite the simplicity of the model concerning internal dynamics it is found to predict sufficiently well the behavior in compact protein structures and it is considered here to be sufficient to elucidate the general trends and guide and motivate the design of relaxation-optimized pulse sequences.

10.2 Experimental details for the measurement and simulation of ubiquitin proton longitudinal relaxation

The experimental proton inversion recovery data in figure 2.14 on page 77 were recorded at 298K on a 600 MHz Varian INOVA spectrometer using an inversion-recovery experiment: either a selective inversion pulse (panel (a), ISNOB5 pulse (Kupce et al. (1995)) centered at 8.5 ppm covering a range of 4 ppm, pulse length of 1.88 ms) or a hard 180° pulse (panel (b)) was followed by a strong magnetic field gradient (ca. 30 G/cm, 1 ms) and a relaxation delay T_{relax} ("recovery time"), after which the longitudinal magnetization of the amide proton spins, or the aliphatic spins, was read out using a ^{1}H-^{15}N (or a ^{1}H-^{13}C) HSQC experiment (Kay et al. (1992)). A long delay (5 seconds) was inserted between two such scans to ensure full recovery to thermal equilibrium.

In the simulations, the overall tumbling correlation time was assumed to be 4 ns, which is in agreement with values reported in the literature (Brutscher et al. (1997); Schneider et al. (1992); Song et al. (2007); Tjandra et al. (1995); Wang et al. (2003)), methyl rotation was included as in equation 10.2 with a correlation time τ_i of 50 ps (see section 10.1).

10.3 Previously proposed experiment for the measurement of amide ^{1}H-^{1}H RDCs in weakly aligned proteins

Figure 10.3 shows the pulse scheme of the 3D SS-HMQC2 experiment proposed by Wu and Bax (2002). Narrow and wide pulses correspond to flip angles of 90° and 180°, respectively. Low amplitude corresponds to selective (1 ms) 90° H_2O pulses. All pulses are along x, unless specified otherwise. The four shaped ^{15}N 180° pulses

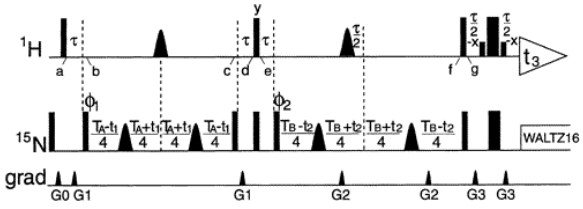


Figure 10.1

are of the hyperbolic secant adiabatic inversion type with durations of 1 ms each; shaped ^{1}H pulses are 180° reBURP (Geen and Freeman (1991). ^{13}C decoupling (not

shown) is used for samples that include ^{13}C enrichment. Prior to time point g, the ^{1}H carrier is at 8 ppm; after g it is switched to H_2O. Delay durations: τ = 5.3 ms; T_A and T_B are typically in the 30-80 ms range. Phase cycling: φ_1 = x,-x; φ_2 = 2x,2(-x); receiver = x,-x,-x,x. Quadrature in both ^{15}N dimensions is obtained by altering φ_1 and φ_2, respectively, in the regular States-TPPI manner. Pulsed field gradients are sine-bell shaped with durations of $G_{0,1,2,3}$ = 1, 1, 0.5, 0.5 ms, with peak amplitudes of 25 G/cm, and directions $G_{0,1,2,3}$ = (xy), x, y, z.

J|A|C|S
COMMUNICATIONS

Published on Web 02/17/2007

Automated Spectral Compression for Fast Multidimensional NMR and Increased Time Resolution in Real-Time NMR Spectroscopy

Ewen Lescop, Paul Schanda, Rodolfo Rasia, and Bernhard Brutscher*

IBS, Institut de Biologie Structurale Jean-Pierre Ebel, 41 rue Jules Horowitz, F-38027 Grenoble; CEA; CNRS; UJF

Received December 14, 2006; E-mail: bernhard.brutscher@ibs.fr

Multidimensional (nD) NMR spectroscopy is a versatile tool for the study of molecular structure, dynamics, and kinetics in solution. In an ideal NMR experiment, each correlation peak reports on the physicochemical properties of the molecular environment at the site of a nuclear spin. For larger molecules, or molecules with a low chemical shift dispersion, an increasing number of spins needs to be correlated in a single 2D, 3D, 4D, or higher-dimensional spectrum to solve the NMR overlap problem. A major drawback of nD NMR is the increase in experimental time by about 2 orders of magnitude for each additional dimension. In protein NMR, the acquisition time of many common experiments exceeds the time required to achieve a sufficient signal-to-noise ratio. To make nD NMR more time-efficient, alternative sampling schemes are required that allow free adjustment of the acquisition time to the intrinsic sensitivity of the experimental setup (sample, spectrometer, pulse sequence). Many elegant approaches for fast data acquisition have been proposed recently, for example, nonlinear data sampling, projection NMR, Hadamard spectroscopy, spatial frequency encoding,[1] and fast-pulsing techniques.[2] Here we present automated spectral compression (ASCOM), a complementary fast NMR technique that exploits available knowledge from prior collected data to increase the peak density in spectral space without loss of information. We demonstrate the use of ASCOM for the study of molecular kinetics based on a series of $^1H-^{15}N$ correlation spectra recorded at 1 s^{-1} rates and for chemical shift assignment using 4D experiments that can be performed in a few hours.

In Fourier-transform (FT) NMR, the choice of the spectral width SW determines the number of (complex) data points n to be recorded to achieve a given spectral resolution $\Delta\nu$ via the relation $n = SW/\Delta\nu$. A reduced spectral width results in fewer repetitions of the pulse sequence and thus in shorter experimental times. If SW is chosen smaller than the chemical shift range of the observed nuclear spins, this results in spectral aliasing (folding). Resonance frequency offsets ν_0 outside the chosen spectral width give rise to NMR signals at $\nu = \epsilon^k\,[\nu_0 - k\ \mathrm{sign}(\nu_0)SW]$ depending whether the time domain data are complex ($\epsilon = 1$) or real ($\epsilon = -1$). The variable k is the number of times the peak has been aliased. To avoid accidental peak overlap, spectral aliasing is generally kept to a minimum. If the chemical shifts of one or several nuclear spin species are already known from previously recorded spectra, this information can be exploited to minimize SW for further data acquisition involving the same nuclear spins. For example, $^1H-^{15}N$ and $^1H-^{13}C$ correlation experiments are generally performed during the early stage of an NMR protein study. We have developed a simple algorithm that optimizes SW_X (X = ^{15}N or ^{13}C) based on the (^1H, X) peak list of a protein (ASCOM-1). Spectra are calculated for each SW_X by representing the cross-peaks by ellipses defined by line width parameters in the two dimensions (Supporting Information, Figure S2). The number of resolved cross-peaks is computed as a function of SW_X. Here we focus on complex data as obtained from STATES,[3a] STATES-TPPI,[3b] or echo/antiecho[3c] quadrature detection. For complex NMR data, the ASCOM optimization is independent of the X carrier frequency and thus requires less computation. Fast computation becomes important when using ASCOM directly on the NMR spectrometer for automated optimization of acquisition parameters. Figure 1a shows that, for small to medium sized proteins, high degrees of compression with minimal or no loss of spectral information are achieved from ASCOM-1 optimization of the ^{15}N spectral width. $^1H-^{15}N$ spectra recorded for ubiquitin using SW_N = 25 ppm (85 resolved peaks) and SW_N = 5.7 ppm (81 resolved peaks) are shown in Figure 1b.

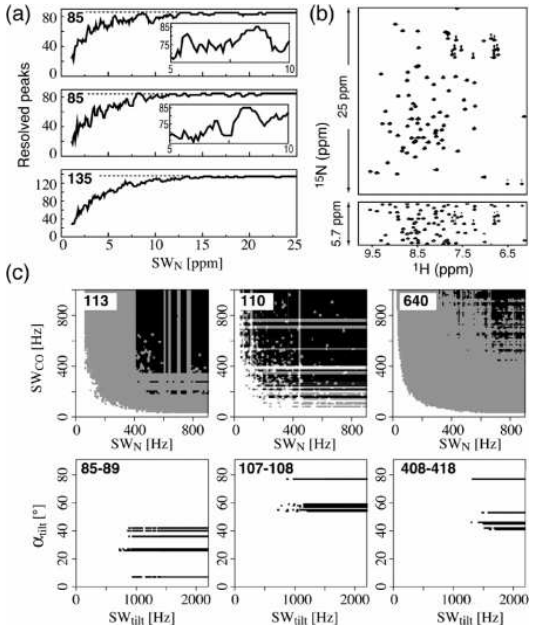


Figure 1. (a) ASCOM-1 optimization of SW_N ($\mathbf{B}_0$ = 18.8 T, $\Delta\nu_H = \Delta\nu_N$ = 30 Hz) for $^1H-^{15}N$ correlation experiments of different proteins: 76-residue ubiquitin (top), a 103-residue fragment of Hyl1 (Hyl1-1)[4] (center), and 147-residue WZB (bottom, BMRB 6934). ASCOM-1 is particularly attractive for small proteins with a low intrinsic peak density. (b) Ubiquitin $^1H-^{15}N$ correlation spectra recorded for different ^{15}N spectral widths. (c) ASCOM-2_{SW} (top) and ASCOM-2_{proj} (bottom) optimizations assuming $\mathbf{B}_0$ = 14.1 T, $\Delta\nu_H$ = 30 Hz, $\Delta\nu_N = \Delta\nu_{CO}$ = 15 Hz. Results are shown for the natively unstructured γ-synuclein (left, BMRB 7244), Hyl1-1 (center), and malate synthase G (right, BMRB 5471). In ASCOM-2_{proj} a peak is counted as resolved if it does not overlap in at least one of the two projections. Parameter combinations for which a maximum of cross-peaks (number given in the upper left corner) are resolved are color-coded in black. The additional gray surface areas indicate the ASCOM-2_{SW} results for a number of resolved cross-peaks equal to the maximum obtained by ASCOM-2_{proj}. For smaller proteins, the two approaches yield comparable results. For large or unstructured protein, only ASCOM-2_{SW} resolves all possible peaks while providing significant data compression.

As a first application of interest, a series of SOFAST-HMQC $^1H-^{15}N$ spectra[2a] can be recorded with a repetition rate of ~1 spectrum/second to follow a kinetic process in real time. So far,

10.1021/ja068949u CCC: $37.00

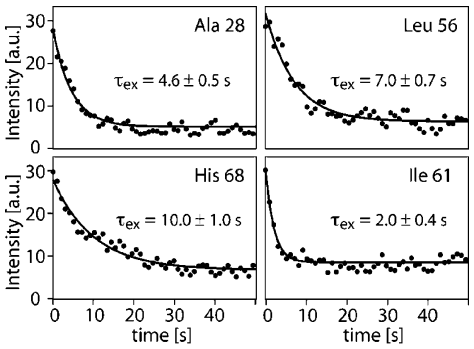


Figure 2. H/D exchange rates measured for amide sites in ubiquitin (1.4 mM, pH 10.8, 2 M urea, 25 °C, 18.8 T). The exchange reaction was initiated by fast injection[2e] of a 50 μL protein−H_2O solution into 350 μL of D_2O. The dead time of the experiment was ∼2 s. The ASCOM-1 optimized ^{15}N spectral width for this sample (48 cross-peaks observable) was 4.9 ppm. SOFAST-HMQC[2a] spectra were recorded with eight complex points in t_1 (5 ms recycle delay) in an overall time of 1 s.

the time resolution of 2D real-time protein NMR has been limited to ∼5 s using standard SOFAST-HMQC. Hadamard-encoded SOFAST-HMQC[2b] or spatially encoded ultraSOFAST HMQC[2c] also yield 1 s^{-1} repetition rates, but the former focuses on a set of selected amide sites, and the latter requires very high protein concentrations. ASCOM-optimized SOFAST-HMQC yields 1 s time resolution simultaneously for all amide sites resolved in the ^{1}H−^{15}N correlation map using moderate protein concentrations. To illustrate this method, we have measured fast amide hydrogen−deuterium exchange in ubiquitin under weak denaturing conditions at high pH. The H/D exchange rates measured under such conditions provide useful information on global unfolding events in the protein.[5] Representative examples of exchange curves are shown in Figure 2. These results clearly demonstrate that real-time 2D NMR measurements of kinetic events down to characteristic time constants τ_{ex} of a few seconds are accessible to ASCOM-1 optimized SOFAST-HMQC. Other interesting applications of ASCOM-1 comprise ^{15}N relaxation measurements, ^{13}C, ^{15}N-edited 3D and 4D ^{1}H−^{1}H NOESY, and 3D H−N−X correlation experiments as required for resonance assignment.

The ASCOM approach can be extended to 3-dimensional spectral space if chemical shift information for three correlated nuclei is available. In ASCOM-2, two spectral parameters are optimized simultaneously, either (i) the spectral widths in the two indirect dimensions (ASCOM-2_{SW}) or (ii) a tilt angle and a tilted spectral width for recording a pair of plane projections (ASCOM-2_{proj}).[6] Here, we focus on correlation experiments required for protein resonance assignment. Recording of 4D NMR spectra with high resolution generally requires unreasonably long acquisition times of >1 week. Therefore, resonance assignment is usually performed based on a set of 3D spectra. With ASCOM-2, however, recording of 4D H−N−CO−X spectra becomes possible in an experimental time comparable to a 3D experiment. The peak density is optimized on the basis of the correlated frequency information obtained from a prior recorded 3D H−N−CO data set. The small number of HNCO peaks (about one per residue), makes ASCOM-2 an efficient tool. For small proteins, ASCOM-2_{proj} yields highest data compression, while still resolving most of the correlation peaks. For larger or unstructured proteins, however, fewer peaks are resolved in the ASCOM-2_{proj}-optimized projections than in the 3D spectrum, as illustrated in Figure 1c.

Even for the 723-residue MSG protein, ASCOM-2_{SW} yields optimized ^{15}N, ^{13}CO spectral widths of only ∼500 Hz.

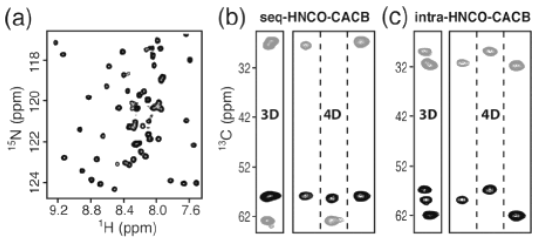


Figure 3. (a) Central part of a ^{1}H−^{15}N HSQC spectrum of Hyl1-1. (b, c) 2D ^{1}H−^{13}C strips extracted from 3D and 4D seq-HNCO-CACB and intra-HNCO-CACB spectra at ω_H = 8.2 ppm, ω_N = 122.2 ppm, and ω_{CO} = 174.6, 175.0, and 179.7 ppm for the 4D data sets. The 4D spectra shown in panels b and c were recorded at 600 MHz on a 0.4 mM sample in an experimental time of 20 and 40 h, respectively, using ASCOM-2_{SW} optimized ^{15}N and ^{13}CO spectral widths of 141 and 483 Hz, respectively.

Figure 3 shows 2D strips extracted from ASCOM-2_{SW} optimized sequential and intraresidue 4D HNCO-CACB experiments (Figure S5) of Hyl1-1. These experiments provide chemical shift information for five nuclei in a single data set. The 4D spectra solve ambiguities remaining in a 3D spectrum because of overlapping ^{1}H−^{15}N correlations, by dispersing the peaks along an additional CO dimension. Therefore they greatly facilitate spectral analysis and sequential assignment, especially for the crowded central spectral regions. By combining ASCOM-2 optimization with the BEST concept,[2d] introduced recently for fast pulsing NMR, a complete 4D HNCO-CACB data set can be recorded in less than 4 h as long as sensitivity is not the limiting factor. This is demonstrated for a 2 mM sample of ubiquitin in Supporting Information Figure S6.

In summary, we have introduced ASCOM a simple tool that optimizes the peak density in 2D or 3D spectral space in order to reduce data acquisition times. We expect that this approach will prove very useful for the study of larger molecular systems, unfolded or partially folded proteins using higher-dimensional NMR, as well as for the study of unidirectional biophysical processes by real-time 2D NMR spectroscopy.

Acknowledgment. We thank J. Boisbouvier (IBS) for making the Hyl1-1 sample available and A. Favier (IBS) for implementing ASCOM on Varian spectrometers.

Supporting Information Available: Further details on the ASCOM algorithm, BEST HNCO-CACB pulse sequences, and ASCOM-optimized 4D spectra of ubiquitin. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Reviewed in Freeman, R.; Kupce, E. *J. Biomol. NMR* **2003**, *27*, 101−113. (b) Malmodulin, D.; Billeter, M. *Prog. Nucl. Magn. Reson. Spectrosc.* **2005**, *46*, 109−129.
(2) (a) Schanda, P.; Brutscher, B. *J. Am. Chem. Soc.* **2005**, *127*, 8014−8015. Schanda, P.; Kupce, E.; Brutscher B. *J. Biomol. NMR* **2005**, *33*, 199−211. (b) Schanda, P.; Brutscher, B. *J. Magn. Reson.* **2006**, *178*, 334−339. (c) Gal, M.; Schanda, P.; Brutscher B.; Frydman, L. *J. Am. Chem. Soc.* **2007**, *129*, 1372−1377. (d) Schanda, P.; Van Melckebeke, H.; Brutscher, B. *J. Am. Chem. Soc.* **2006**, *128*, 9042−9043. (e) Mok, K. H.; Nagashima, T.; Day, I. J.; Jones, J. A.; Jones, C. J. V.; Dobson, C. V.; Hore, P. J. *J. Am. Chem. Soc.* **2003**, *125*, 12484−12492.
(3) (a) States, D. J.; Haberkorn, R. A.; Ruben, D. J. *J. Magn. Reson.* **1982**, *48*, 286−292. (b) Marion, D.; Ikura, M.; Tschudin, R.; Bax, A. *J. Magn. Reson.* **1989**, *85*, 393−399. (c) Palmer, A. G.; Cavanagh, J.; Wright, P. E.; Rance, M. *J. Magn. Reson.* **1994**, *93*, 151−170.
(4) Han, M. H.; Goud, S.; Song, L.; Fedoroff, N. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 1093−1098.
(5) Dempsey, C. E. *Prog. Nucl. Magn. Reson. Spectrosc.* **2001**, *39*, 135−170. Ferraro, D. M.; Lazo, N. D.; Robertson, A. *Biochemistry* **2004**, *43*, 587−594.
(6) Tugarinov, V.; Choy, W.-Y.; Kupce, E.; Kay, L. E. *J. Biomol. NMR* **2004**, *30*, 347−352.

JA068949U

List of scientific publications

Schanda, P. and Brutscher, B. (2005). Very fast two dimensional NMR spectroscopy for real-time investigation of dynamics events in proteins on the time scale of seconds. *J. Am. Chem. Soc.*,127(22): 8014-5.

Schanda, P., Kupce, E. and Brutscher, B. (2005). SOFAST-HMQC experiments for recording of two-dimensional heteronuclear correlation spectra of proteins within a few seconds. *J. Biomol. NMR*, 33(4): 199-211.

Schanda, P. and Brutscher, B. (2006). Hadamard frequency-encoded SOFAST HMQC for ultrafast two-dimensional protein NMR. *J. Magn. Reson*, 178(2): 334-9.

Schanda,P., Forge,V. and Brutscher,B.(2006).HET-SOFAST NMR for fast detection of structural compactness and heterogeneity along polypeptide chains. *Magn. Reson. Chem*, 44: S177-84.

Schanda, P. Van Melckebeke, H. and Brutscher, B. (2006). Speeding up three-dimensional protein NMR experiments to a few minutes. *J. Am. Chem. Soc.*, 128(28): 9042-3.

Gal, M., Schanda, P., Brutscher, B. and Frydman, L. (2007). UltraSOFAST HMQC NMR and the Repetitive Acquisition of 2D Protein Spectra at Hertz Rates. *J. Am. Chem. Soc.*, 129(5): 1372-7.

Schanda, P., Lescop, E., Falge, M., Sounier, R., Boisbouvier, J. and Brutscher, B. (2007). Sensitivity-optimized experiment for the measurement of residual dipolar couplings between amide protons. *J. Biomol. NMR*, 38(1): 47-55.

Lescop, E., Schanda, P., Rasia, R. and Brutscher, B. (2007). Automated spectral compression for fast multidimensional NMR and increased time-resolution in real-time NMR. *J. Am. Chem. Soc.*, 129(10): 2756-7. (see preceding pages)

Lescop, E., Schanda, P. and Brutscher, B. (2007). A set of BEST triple-resonance experiments for time-optimized protein resonance assignment. *J. Magn. Reson.*, 187(1): 163-9.

Schanda, P., Forge, V. and Brutscher, B. (2007). Protein folding and unfolding studied at atomic resolution by fast two-dimensional NMR spectroscopy. *Proc. Natl. Acad. Sci. USA*, 104(27): 11257-62.

Bibliography

Abragam, A. (1961). *Principles of Nuclear Magnetism*. Clarendon Press, Oxford.

Akasaka, K., Ishima, R., and Shibata, S. (1990). Proton spin relaxation in biopolymers at high magnetic fields. *Physica B*, 164(1-2):163–179.

Albert, M., Cates, G., Driehuys, B., Happer, W., Saam, B., Springer, C., and Wishnia, A. (1994). Biological magnetic resonance imaging using laser-polarized xe. *Nature*, 370:199–201.

Andersson, P., Weigelt, J., and Otting, G. (1998). Spin-state selection filters for the measurement of heteronuclear one-bond coupling constants. *J Biomol NMR*, 12(3):435–41.

Arai, M. and Kuwajima, K. (2000). Role of the molten globule state in protein folding. *Adv Protein Chem*, 53(0065-3233 (Print)):209–82.

Arrington, C. and Robertson, A. (1997). Microsecond protein folding kinetics from native-state hydrogen exchange. *Biochemistry*, 36:8686–8691.

Atreya, H., Eletsky, A., and Szyperski, T. (2005). Resonance assignment of proteins with high shift degeneracy based on 5d spectral information encoded in g2ft nmr experiments. *J Am Chem Soc*, 127:4554–4555.

Atreya, H. S. and Szyperski, T. (2004). G-matrix fourier transform nmr spectroscopy for complete protein resonance assignment. *Proc Natl Acad Sci U S A*, 101(26):9642–7.

Aue, W., Bartholdi, E., and Ernst, R. (1976). Two-dimensional spectroscopy. application to nuclear magnetic resonance. *J. Chem. Phys.*, 64:2229.

Babu, C. R., Hilser, V. J., and Wand, A. J. (2004). Direct access to the cooperative substructure of proteins and the protein ensemble via cold denaturation. *Nat Struct Mol Biol*, 11(4):352–7.

Balbach, J. (2000). Compaction during protein folding studied by real-time nmr diffusion experiments. *J. Am. Chem. Soc.*, 122:5887–5888.

Balbach, J., Forge, V., Lau, W. S., van Nuland, N. A., Brew, K., and Dobson, C. M. (1996). Protein folding monitored at individual residues during a two-dimensional nmr experiment. *Science*, 274(5290):1161–3.

Balbach, J., Forge, V., Van Nuland, N. A., Winder, S. L., Hore, P. J., and Dobson, C. M. (1995). Following protein folding in real time using nmr spectroscopy. *Nat. Struct. Biol.*, 2:865–870.

Barkhuijsen, H., De Beer, R., Bovee, W., and Van Ormondt, D. (1985). Retrieval of frequencies, amplitudes, damping factors and phases from time-domain signals using a linear least-squares procedure. *J. Magn. Reson.*, 61(3):465–481.

Bax, A. and Pochapsky, S. (1992). Optimized recording of heteronuclear multidimenisional nmr spectra using pulsed field gradients. *J. Magn. Reson.*, 99(3):638–643.

Bax, A., Vuister, G. W., Grzesiek, S., Delaglio, F., Wang, A. C., Tschudin, R., and Zhu, G. (1994). Measurement of homo- and heteronuclear j couplings from quantitative j correlation. *Methods Enzymol*, 239(0076-6879 (Print)):79–105.

Beraud, S., Bersch, B., Brutscher, B., Gans, P., Barras, F., and Blackledge, M. (2002). Direct structure determination using residual dipolar couplings: reaction-site conformation of methionine sulfoxide reductase in solution. *J Am Chem Soc*, 124(46):13709–15.

Bernado, P., Blanchard, L., Timmins, P., Marion, D., Ruigrok, R. W. H., and Blackledge, M. (2005). A structural model for unfolded proteins from residual dipolar couplings and small-angle x-ray scattering. *Proc Natl Acad Sci U S A*, 102(47):17002–7.

Bersch, B., Rossy, E., Coves, J., and Brutscher, B. (2003). Optimized set of two-dimensional experiments for fast sequential assignment, secondary structure determination, and backbone fold validation of 13c/15n-labelled proteins. *J Biomol NMR*, 27(1):57–67.

Bhattacharyya, R. and Frydman, L. (2006). Ultrafast solid-state 2d nmr experiments via orientational encoding. *J Am Chem Soc*, 128(50):16014–16015.

Blackledge, M. (2005). Recent progress in the study of biomolecular structure and dynamics in solution from residual dipolar couplings. *Prog. Nucl. Magn. Reson. Spectrosc.*, 46:23–61.

Blackledge, M. J., Medvedeva, S., Poncin, M., Guerlesquin, F., Bruschi, M., and Marion, D. (1995). Structure and dynamics of ferrocytochrome c553 from desulfovibrio vulgaris studied by nmr spectroscopy and restrained molecular dynamics. *J Mol Biol*, 245(5):661–81.

Boisbouvier, J., Delaglio, F., and Bax, A. (2003). Direct observation of dipolar couplings between distant protons in weakly aligned nucleic acids. *Proc Natl Acad Sci U S A*, 100(20):11333–8.

Bougault, C., Feng, L., Glushka, J., Kupce, E., and Prestegard, J. H. (2004). Quantitation of rapid proton-deuteron amide exchange using hadamard spectroscopy. *J Biomol NMR*, 28(4):385–90.

Bouvignies, G., Bernado, P., Meier, S., Cho, K., Grzesiek, S., Bruschweiler, R., and Blackledge, M. (2005). Identification of slow correlated motions in proteins using residual dipolar and hydrogen-bond scalar couplings. *Proc Natl Acad Sci U S A*, 102(39):13885–90.

Bouvignies, G., Meier, S., Grzesiek, S., and Blackledge, M. (2006). Ultrahigh-resolution backbone structure of perdeuterated protein gb1 using residual dipolar couplings from two alignment media. *Angew Chem Int Ed Engl*, 45(48):8166–9.

Bowers, C. and Weitekamp, D. (1986). Transformation of symmetrization order in nuclear spin magnetization by chemical reaction and nuclear magnetic resonance. *Phys Rev Lett*, 57:2645–48.

Browne, D. T., Kenyon, G. L., Packer, E. L., Sternlicht, H., and Wilson, D. M. (1973). Studies of macromolecular structure by carbon-13 nuclear magnetic resonance. ii. specific labeling approach to the study of histidine residues in proteins. *J. Am. Chem. Soc.*, 95(4):1316–1323.

Bruschweiler, R. and Zhang, F. (2004). Covariance nuclear magnetic resonance spectroscopy. *J Chem Phys*, 120(11):5253–60.

Brutscher, B. (2000). Principles and applications of cross-correlated relaxation in biomolecules. *Concepts Magnetic Reson.*, 12:207–229.

Brutscher, B. (2002). Intraresidue hnca and cohnca experiments for protein backbone resonance assignment. *J Magn Reson*, 156(1):155–9.

Brutscher, B. (2004). Combined frequency- and time-domain nmr spectroscopy. application to fast protein resonance assignment. *J Biomol NMR*, 29(1):57–64.

Brutscher, B., Bruschweiler, R., and Ernst, R. R. (1997). Backbone dynamics and structural characterization of the partially folded a state of ubiquitin by 1h, 13c, and 15n nuclear magnetic resonance spectroscopy. *Biochemistry*, 36(42):13043–53.

Brutscher, B., Cordier, F., Simorre, J., Caffrey, M., and Marion, D. (1995a). High-resolution 3d hncoca experiment applied to a 28 kda paramagnetic protein. *J Biomol NMR*, 5:202–206.

Brutscher, B., Morelle, N., Cordier, F., and Marion, D. (1995b). Determination of an initial set of noe-derived distance constraints for the structure determination of 15n,13c labeled proteins. *J Magn Reson B*, 109(2):238–242.

Brutscher, B., Simorre, J., Caffrey, M., and Marion, D. (1994). Design of a complete set of two-dimensional triple-resonance experiments for assigning labeled proteins. *J. Magn. Reson. B*, 105:77–82.

Buevich, A. V., Dai, Q. H., Liu, X., Brodsky, B., and Baum, J. (2000). Site-specific nmr monitoring of cis-trans isomerization in the folding of the proline-rich collagen triple helix. *Biochemistry*, 39(15):4299–308.

Bushmarina, N. A., Blanchet, C. E., Vernier, G., and Forge, V. (2006). Cofactor effects on the protein folding reaction: acceleration of alpha-lactalbumin refolding by metal ions. *Protein Sci*, 15(4):659–71.

Cai, S., Seu, C., Kovacs, Z., Sherry, A. D., and Chen, Y. (2006). Sensitivity enhancement of multidimensional nmr experiments by paramagnetic relaxation effects. *J Am Chem Soc*, 128(41):13474–8.

Caravan, P., Ellison, J., McMurry, T., and Lauffer, R. (1999). Gadolinium(iii) chelates as mri contrast agents: Structure, dynamics, and applications. *Chem. Rev.*, 99(9):2293–2352.

Cavanagh, J., Fairbrother, W., Palmer, A., and Skelton, N. (1995). *Protein NMR spectroscopy*. Academic Press.

Cavanagh, J., Palmer, A. G. I., Wright, P. E., and Rance, M. (1991). Sensitivity improvement in proton-detected 2-dimensional heteronuclear relay spectroscopy. *J. Magn. Reson.*, 91:429–436.

Cavanagh, J. and Rance, M. (1990). Sensitivity improvements in isotropic mixing experiments. *J. Magn. Reson.*, 88:72–85.

Champier, L., Sibille, N., Bersch, B., Brutscher, B., Blackledge, M., and Coves, J. (2002). Reactivity, secondary structure, and molecular topology of the escherichia coli sulfite reductase flavodoxin-like domain. *Biochemistry*, 41(11):3770–80.

Chen, J., De Angelis, A., Mandelshtam, V., and Shaka, A. (2003). Progress on the two-dimensional filter diagonalization method. an efficient doubling scheme for two-dimensional constant-time nmr. *J Magn Reson*, 162(1):74–89.

Chrysina, E. D., Brew, K., and Acharya, K. R. (2000). Crystal structures of apo- and holo-bovine alpha-lactalbumin at 2. 2-a resolution reveal an effect of calcium on inter-lobe interactions. *J Biol Chem*, 275(47):37021–9.

Chung, H. S., Khalil, M., Smith, A. W., Ganim, Z., and Tokmakoff, A. (2005). Conformational changes during the nanosecond-to-millisecond unfolding of ubiquitin. *Proc Natl Acad Sci U S A*, 102(3):612–7.

Coggins, B., Venters, R., and Zhou, P. (2005). Filtered backprojection for the reconstruction of a high-resolution (4,2)d ch3-nhnoesy spectrum on a 29 kda protein. *J. Am. Chem. Soc.*, 127(33):11562–11563.

Cordier, F. and Grzesiek, S. (2002). Temperature-dependence of protein hydrogen bond properties as studied by high-resolution nmr. *J Mol Biol*, 317(5):739–52.

Cordier, F. and Grzesiek, S. (2004). Quantitative comparison of the hydrogen bond network of a-state and native ubiquitin by hydrogen bond scalar couplings. *Biochemistry*, 43(35):11295–301.

Cornilescu, G., Marquardt, J. L., Ottiger, M., and Bax, A. (1998). Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. *J. Am. Chem. Soc.*, 120:6836–6837.

Delaglio, F., Grzesiek, S., Vuister, G., Zhu, G., Pfeifer, J., and Bax, A. (1995). Nmrpipe: a multidimensional spectral processing system based on unix pipes. *J. Biomol. NMR.*, 6:277–293.

Dempsey, C. (2001). Hydrogen exchange in peptides and proteins using nmr spectroscopy. *Prog. NMR Spectr.*, 39:135–170.

Deschamps, M. and Campbell, I. D. (2006). Cooling overall spin temperature: protein nmr experiments optimized for longitudinal relaxation effects. *J Magn Reson*, 178(2):206–11.

Diercks, T., Daniels, M., and Kaptein, R. (2005). Extended flip-back schemes for sensitivity enhancement in multidimensional hsqc-type out-and-back experiments. *J Biomol NMR*, 33(4):243–59.

Dill, K. A., Fiebig, K. M., and Chan, H. S. (1993). Cooperativity in protein-folding kinetics. *Proc Natl Acad Sci U S A*, 90(5):1942–6.

Ding, K. and Gronenborn, A. (2002). Novel 2d triple-resonance nmr experiments for sequential resonance assignments of proteins. *J. Magn. Reson.*, 156:262–268.

Dobson, C. and Hore, P. (1998). Kinetic studies of protein folding using nmr spectroscopy. *Nat. Struct. Biol.*, 5:504–507.

Dobson, C. M. (2004). Principles of protein folding, misfolding and aggregation. *Semin Cell Dev Biol*, 15(1):3–16.

Dunker, A. K. and Obradovic, Z. (2001). The protein trinity–linking function and disorder. *Nat Biotechnol*, 19(9):805–6.

Ecker, D. J., Butt, T. R., Marsh, J., Sternberg, E., Shatzman, A., Dixon, J. S., Weber, P. L., and Crooke, S. T. (1989). Ubiquitin function studied by disulfide engineering. *J Biol Chem*, 264(3):1887–93.

Eisenmesser, E. Z., Bosco, D. A., Akke, M., and Kern, D. (2002). Enzyme dynamics during catalysis. *Science*, 295(5559):1520–3.

Eletsky, A., Atreya, H., Liu, G., and Szyperski, T. (2005). Probing structure and functional dynamics of (large) proteins with aromatic rings: L-gft-trosy (4,3)d hcch nmr spectroscopy. *J. Am. Chem. Soc.*, 127(42):14578–14579.

Eletsky, A., Moreira, O., Kovacs, H., and Pervushin, K. (2003). A novel strategy for the assignment of side-chain resonances in completely deuterated large proteins using 13c spectroscopy. *J Biomol NMR*, 26(2):167–79.

Emsley, L. and Bodenhausen, G. (1992). Optimization of shaped selective pulses for nmr using a quaternion description of their overall propagators. *J. Magn. Reson.*, 97:135–148.

Ernst, R. and Anderson, W. (1966). Application of fourier transform spectroscopy to magnetic resonance. *Rev. Sci. Instrum.*, 37:93.

Ernst, R., Bodenhausen, G., and A., W. (1987). *Principles of Nuclear Magnetic Resonances in One and Two Dimensions*. Oxford Science Publications.

Feliz, M., Garcia, J., Aragon, E., and Pons, M. (2006). Fast 2d nmr ligand screening using hadamard spectroscopy. *J. Am. Chem. Soc.*, 128(22):7146–7147.

Ferraro, D. M., Lazo, N. D., and Robertson, A. D. (2004). Ex1 hydrogen exchange and protein folding. *Biochemistry*, 43(3):587–94.

Fiala, R., Czernek, J., and Sklenar, V. (2000). Transverse relaxation optimized triple-resonance nmr experiments for nucleic acids. *J. Biomol. NMR*, 16(4):291–302.

Fitzgerald, R. J. and Swaisgood, H. E. (1989). Binding of ions and hydrophobic probes to alpha-lactalbumin and kappa-casein as determined by analytical affinity chromatography. *Arch Biochem Biophys*, 268(1):239–48.

Forge, V., Wijesinha, R. T., Balbach, J., Brew, K., Robinson, C. V., Redfield, C., and Dobson, C. M. (1999). Rapid collapse and slow structural reorganisation during the refolding of bovine alpha-lactalbumin. *J Mol Biol*, 288(4):673–88.

Frauenfelder, H., Sligar, S. G., and Wolynes, P. G. (1991). The energy landscapes and motions of proteins. *Science*, 254(5038):1598–603.

Freeman, R. and Kupce, E. (2003). New methods for fast multidimensional nmr. *J Biomol NMR*, 27(2):101–13.

Frydman, L. and Blazina, D. (2007). Ultrafast two-dimensional nuclear magnetic resonance spectroscopy of hyperpolarized solutions. *Nature Physics*, 3:415–419.

Frydman, L., Lupulescu, A., and Scherf, T. (2003). Principles and features of single-scan two-dimensional nmr spectroscopy. *J Am Chem Soc*, 125(30):9204–17.

Frydman, L., Scherf, T., and Lupulescu, A. (2002). The acquisition of multidimensional nmr spectra within a single scan. *Proc Natl Acad Sci U S A*, 99(25):15858–62.

Gal, M., Mishkovsky, M., and Frydman, L. (2006). Real-time monitoring of chemical transformations by ultrafast 2d nmr spectroscopy. *J Am Chem Soc*, 128(3):951–6.

Gal, M., Schanda, P., Brutscher, B., and Frydman, L. (2007). Ultrasofast hmqc nmr and the repetitive acquisition of 2d protein spectra at hz rates. *J Am Chem Soc*, 129(5):1372–7.

Gardner, K. H. and Kay, L. E. (1998). The use of 2h, 13c, 15n multidimensional nmr to study the structure and dynamics of proteins. *Annu Rev Biophys Biomol Struct*, 27(1056-8700 (Print)):357–406.

Geen, H. and Freeman, R. (1991). Band-selective radiofrequency pulses. *J. Magn. Reson.*, 93:93–141.

Grobler, J. A., Wang, M., Pike, A. C., and Brew, K. (1994). Study by mutagenesis of the roles of two aromatic clusters of alpha-lactalbumin in aspects of its action in the lactose synthase system. *J Biol Chem*, 269(7):5106–14.

Grzesiek, S. and Bax, A. (1993). The importance of not saturating h2o in protein nmr - application to sensitivity enhancement and noe measurements. *J. Am. Chem. Soc.*, 115(26):12593–12594.

Gustafsson, L., Leijonhufvud, I., Aronsson, A., Mossberg, A.-K., and Svanborg, C. (2004). Treatment of skin papillomas with topical alpha-lactalbumin-oleic acid. *N Engl J Med*, 350(26):2663–72.

Hadamard, J. (1893). Resolution d'une question relative aux determinants. *Bull. Sci. Math.*, 17:240–248.

Hajduk, P., Augeri, D., Mack, J., Mendoza, R., Yang, J., Betz, S., and Fesik, S. (2000). Nmr-based screening of proteins containing c-13-labeled methyl groups. *J. Am. Chem. Soc.*, 122:7898–7904.

Hiller, S., Fiorito, F., Wuthrich, K., and Wider, G. (2005a). Automated projection spectroscopy (apsy). *Proc Natl Acad Sci U S A*, 102(31):10876–81.

Hiller, S., Wider, G., Etezady-Esfarjani, T., Horst, R., and Wuthrich, K. (2005b). Managing the solvent water polarization to obtain improved nmr spectra of large molecular structures. *J Biomol NMR*, 32(1):61–70.

Hoch, J. and Stern, A. (1996). *NMR Data processing*. Wiley-Liss, New York.

Hoch, J. and Stern, A. (2001). Maximum entropy reconstruction, spectrum analysis and deconvolution in multidimensional magnetic resonance. *NMR Biol. Macromol. Pt. A*, 338:159.

Hoffmann, B., Eichmuller, C., Steinhauser, O., and Konrat, R. (2005). Rapid assessment of protein structural stability and fold validation via nmr. *Methods Enzymol*, 394(0076-6879 (Print)):142–75.

Hus, J. C., Marion, D., and Blackledge, M. (2001). Determination of protein backbone structure using only residual dipolar couplings. *J Am Chem Soc*, 123(7):1541–2.

Huyghues-Despointes, B. M., Scholtz, J. M., and Pace, C. N. (1999). Protein conformational stabilities can be determined from hydrogen exchange rates. *Nat Struct Biol*, 6(10):910–2.

Hwang, T., Mori, S., Shaka, A., and Van Zijl, P. (1997). Application of phase-modulated clean chemical exchange spectroscopy (cleanex-pm) to detect water-protein proton exchange and intermolecular noes. *J. Am. Chem. Soc.*, 119:6203–6204.

Ishikawa, N., Chiba, T., Chen, L. T., Shimizu, A., Ikeguchi, M., and Sugai, S. (1998). Remarkable destabilization of recombinant alpha-lactalbumin by an extraneous n-terminal methionyl residue. *Protein Eng*, 11(5):333–5.

Jackson, S. E. (2006). Ubiquitin: a small protein folding paradigm. *Org Biomol Chem*, 4(10):1845–53.

Jeannerat, D. (2007). Computer optimized spectral aliasing in the indirect dimension of (1)h-(13)c heteronuclear 2d nmr experiments. a new algorithm and examples of applications to small molecules. *J. Magn. Reson.*, 186(1):112–122.

Jeener, J. (1971). Ampere international summer school. In *Basko Polje, Yugoslavia, September*.

Kataoka, M., Kuwajima, K., Tokunaga, F., and Goto, Y. (1997). Structural characterization of the molten globule of alpha-lactalbumin by solution x-ray scattering. *Protein Sci*, 6(2):422–30.

Kay, L., Keifer, P., and Saarinen, T. (1992). Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc.*, 114:10663–10665.

Kay, L., Xu, G., and Yamazaki, T. (1994). Enhanced-sensitivity triple-resonance spectroscopy with minimal h2o saturation. *J. Magn. Reson.*, 109:129–133.

Keppetipola, S., Kudlicki, W., Nguyen, B., Meng, X., Donovan, K., and Shaka, A. (2006). From gene to hsqc in under five hours: High-throughput nmr proteomics. *J Am Chem Soc*, 128(14):4508–4509.

Kim, S. and Szyperski, T. (2003). Gft nmr, a new approach to rapidly obtain precise high-dimensional nmr spectral information. *J Am Chem Soc*, 125(5):1385–93.

Kitahara, R. and Akasaka, K. (2003). Close identity of a pressure-stabilized intermediate with a kinetic intermediate in protein folding. *Proc Natl Acad Sci U S A*, 100(6):3167–72.

Kontaxis, G., Delaglio, F., and Bax, A. (2005). Molecular fragment replacement approach to protein structure determination by chemical shift and dipolar homology database mining. *Methods Enzymol*, 394(0076-6879 (Print)):42–78.

Kontaxis, G., Stonehouse, J., Laue, E., and Keeler, J. (1994). The sensitivity of experiments which use gradient pulses for coherence-pathway selection. *J. Magn. Reson. A*, 111:70–76.

Kovacs, H., Moskau, D., and Spraul, M. (2005). Cryogenically cooled probes - a leap in nmr technology. *Prog. NMR Spectr.*, 46:131–155.

Krishna, M. M. G., Hoang, L., Lin, Y., and Englander, S. W. (2004). Hydrogen exchange methods to study protein folding. *Methods*, 34(1):51–64.

Kupce, E., Boyd, J., and Campbell, I. D. (1995). Short selective pulses for biochemical applications. *J Magn Reson B*, 106(3):300–3.

Kupce, E. and Freeman, R. (1993). Techniques for multisite excitation. *J. Magn. Reson. A*, 105:234–238.

Kupce, E. and Freeman, R. (1994). Wideband excitation with polychromatic pulses. *J. Magn. Reson. A*, 108:268–273.

Kupce, E. and Freeman, R. (1995). Adiabatic pulses for wideband inversion and broadband decoupling. *J. Magn. Reson. A*, 115:273–276.

Kupce, E. and Freeman, R. (1996). Optimized adiabatic pulses for wideband spin inversion. *J. Magn. Reson. A*, 118:299–303.

Kupce, E. and Freeman, R. (2003a). Fast multi-dimensional hadamard spectroscopy. *J. Magn. Reson.*, 163:56–63.

Kupce, E. and Freeman, R. (2003b). Projection-reconstruction of three-dimensional nmr spectra. *J. Am. Chem. Soc.*, 125(46):13958–13959.

Kupce, E. and Freeman, R. (2004a). Projection-reconstruction technique for speeding up multidimensional nmr spectroscopy. *J Am Chem Soc*, 126(20):6429–40.

Kupce, E. and Freeman, R. (2004b). The radon transform: A new scheme for fast multidimensional nmr. *Concepts in Magn. Reson.*, 22A(1):4–11.

Kupce, E. and Freeman, R. (2007). Fast multidimensional nmr by polarization sharing. *Magn Reson Chem*, 45(1):2–4.

Kupce, E., Nishida, T., and Freeman, R. (2003). Hadamard nmr spectroscopy. *Prog. Nucl. Magn. Reson. Spectrosc.*, 42:95–122.

Kupce, E. and Wagner, G. (1995). Wideband homonuclear decoupling in protein spectra. *J. Magn. Reson. B*, 109:329–333.

Kutyshenko, V. P. and Cortijo, M. (2000). Water-protein interactions in the molten-globule state of carbonic anhydrase b: an nmr spin-diffusion study. *Protein Sci*, 9(8):1540–7.

Lakomek, N. A., Fares, C., Becker, S., Carlomagno, T., Meiler, J., and Griesinger, C. (2005). Side-chain orientation and hydrogen-bonding imprint supra-tau(c) motion on the protein backbone of ubiquitin. *Angew Chem Int Ed Engl*, 44(47):7776–8.

LeMaster, D. M. (1989). Deuteration in protein proton magnetic resonance. *Methods Enzymol*, 177(0076-6879 (Print)):23–43.

LeMaster, D. M. (1990a). Deuterium labelling in nmr structural analysis of larger proteins. *Q Rev Biophys*, 23(2):133–74.

LeMaster, D. M. (1990b). Uniform and selective deuteration in two-dimensional nmr of proteins. *Annu. Rev. Biophys. Biophys. Chem.*, 19(0883-9182 (Print)):243–66.

Lescop, E., Schanda, P., Rasia, R., and Brutscher, B. (2007). Automated spectral compression for fast multidimensional nmr and increased time resolution in real-time nmr spectroscopy. *J Am Chem Soc*, 129(10):2756–7.

Louhivuori, M., Paakkonen, K., Fredriksson, K., Permi, P., Lounila, J., and Annila, A. (2003). On the origin of residual dipolar couplings from denatured proteins. *J Am Chem Soc*, 125(50):15647–50.

Malmodin, D. and Billeter, M. (2005a). High-throughput analysis of protein nmr spectra. *Prog. NMR Spectrosc.*, 46:109–129.

Malmodin, D. and Billeter, M. (2005b). Signal identification in nmr spectra with coupled evolution periods. *J Magn Reson*, 176:47–53.

Mandelshtam, V. A. (2000). The multidimensional filter diagonalization method. *J Magn Reson*, 144(2):343–56.

Marion, D. (2005). Fast acquisition of nmr spectra using fourier transform of non-equispaced data. *J Biomol NMR*, 32(2):141–50.

Marion, D. (2006). Processing of nd nmr spectra sampled in polar coordinates: a simple fourier transform instead of a reconstruction. *J Biomol NMR*, 36:45–65.

Marion, D., Ikura, M., Tschudin, R., and Bax, A. (1989). Rapid recording of 2d nmr-spectra without phase cycling - application to the study of hydrogen-exchange in proteins. *J. Magn. Reson.*, 85:393–399.

Marion, D. and Wuthrich, K. (1983). Application of phase sensitive two-dimensional correlated spectroscopy (cosy) for measurements of 1h-1h spin-spin coupling constants in proteins. *Biochem. Biophys. Res. Commun.*, 113(3):967–74.

Matsui, S., Sekihara, K., and Kohno, H. (1985). High-speed spatially resolved nmr spectroscopy using phase-modulated spin-echo trains. expansion of the spectral bandwidth by combined use of delayed spin-echo trains. *J. Magn. Reson.*, 64:167–171.

McKnight, C. J., Matsudaira, P. T., and Kim, P. S. (1997). Nmr structure of the 35-residue villin headpiece subdomain. *Nat Struct Biol*, 4(3):180–4.

Meiler, J., Prompers, J. J., Peti, W., Griesinger, C., and Bruschweiler, R. (2001). Model-free approach to the dynamic interpretation of residual dipolar couplings in globular proteins. *J Am Chem Soc*, 123(25):6098–107.

Mergeay, M., Monchy, S., Vallaeys, T., Auquier, V., Benotmane, A., Bertin, P., Taghavi, S., Dunn, J., van der Lelie, D., and Wattiez, R. (2003). Ralstonia metallidurans, a bacterium specifically adapted to toxic metals: towards a catalogue of metal-responsive genes. *FEMS Microbiol Rev*, 27(2-3):385–410.

Mittermaier, A. and Kay, L. E. (2006). New tools provide new insights in nmr studies of protein dynamics. *Science*, 312(5771):224–8.

Mok, K. H., Nagashima, T., Day, I. J., Hore, P. J., and Dobson, C. M. (2005). Multiple subsets of side-chain packing in partially folded states of alpha-lactalbumins. *Proc Natl Acad Sci U S A*, 102(25):8899–904.

Mok, K. H., Nagashima, T., Day, I. J., Jones, J. A., Jones, C. J. V., Dobson, C. M., and Hore, P. J. (2003). Rapid sample-mixing technique for transient nmr and photo-cidnp spectroscopy: applications to real-time protein folding. *J Am Chem Soc*, 125(41):12484–92.

Morelle, N., Brutscher, B., Simorre, J., and Marion, D. (1995). Computer assignment of the backbone resonances of labelled proteins using two-dimensional correlation experiments. *J Biomol NMR*, 5:154–160.

Mulder, F., Mittermaier, A., Hon, B., Dahlquist, F., and Kay, L. (2001). Studying excited states of proteins by nmr spectroscopy. *Nat Struct Biol*, 8(11):932–935.

Muller-Warmuth, W. and Meise-Gresch, K. (1983). Molecular motions and interactions as studied by dynamics nuclear polarization in free radical solutions. *Adv Magn Reson*, 11:1–45.

Nagayama, K., Bachmann, P., Wuthrich, K., and Ernst, R. (1978). The use of cross-sections and of projections in two-dimensional nmr spectroscopy. *J. Magn. Reson.*, 31:133–138.

Neudecker, P., Zarrine-Afsar, A., Choy, W.-Y., Muhandiram, D. R., Davidson, A. R., and Kay, L. E. (2006). Identification of a collapsed intermediate with non-native long-range interactions on the folding pathway of a pair of fyn sh3 domain mutants by nmr relaxation dispersion spectroscopy. *J Mol Biol*, 363(5):958–76.

Nielsen, N., Bildsoe, H., Jakobsen, H., and Sorensen, O. (1989). Composite refocusing sequences and their application for sensitivity enhancement and multiphcity filtration in inept and 2d correlation spectroscopy. *J Magn Reson*, 85:359–380.

Nietlispach, D. (2004). A selective intra-hn(ca)co experiment for the backbone assignment of deuterated proteins. *J Biomol NMR*, 28(2):131–6.

Olejniczak, E. T. and Weiss, M. A. (1990). Are methyl groups relaxation sinks in small proteins ? *J. Magn. Reson.*, 86:148–155.

Orekhov, V., Ibraghimov, I., and Billeter, M. (2001). Munin: A new approach to multi-dimensional nmr spectra interpretation. *J Biomol NMR*, 20:49–60.

Orekhov, V., Ibraghimov, I., and Billeter, M. (2003). Optimizing resolution in multidimensional nmr by three-way decomposition. *J Biomol NMR*, 27:166–173.

Ottiger, M., Delaglio, F., and Bax, A. (1998). Measurement of j and dipolar couplings from simplified two-dimensional nmr spectra. *J Magn Reson*, 131(2):373–8.

Otting, G., Liepinsh, E., and Wüthrich, K. (1991). Protein hydration in aqueous solution. *Science*, 254:974–980.

Palmer, A. G. I., Cavanagh, J., Wright, P. E., and Rance, M. (1991). Sensitivity improvement in proton-detected 2-dimensional heteronuclear correlation nmr spectroscopy. *J. Magn. Reson.*, 93(1):151–170.

Palmer, A. G. r. (2004). Nmr characterization of the dynamics of biomacromolecules. *Chem Rev*, 104(8):3623–40.

Pan, Y. and Briggs, M. S. (1992). Hydrogen exchange in native and alcohol forms of ubiquitin. *Biochemistry*, 31(46):11405–12.

Pannetier, N., Houben, K., Blanchard, L., and Marion, D. (2007). Optimized 3d-nmr sampling for resonance assignment of partially unfolded proteins. *J Magn Reson*, 186(1):142–149.

Pelupessy, P. (2003). Adiabatic single scan two-dimensional nmr spectrocopy. *J Am Chem Soc*, 125(40):12345–50.

Pervushin, K., Riek, R., Wider, G., and Wüthrich, K. (1997). Attenuated t2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to nmr structures of very large biological macromolecules in solution. *Proc. Natl. Acad. Sci. U.S.A.*, 94(23):12366–12371.

Pervushin, K., Vogeli, B., and Eletsky, A. (2002). Longitudinal 1h relaxation optimization in trosy nmr spectroscopy. *J. Am. Chem. Soc.*, 124(43):12898–12902.

Peti, W., Smith, L. J., Redfield, C., and Schwalbe, H. (2001). Chemical shifts in denatured proteins: resonance assignments for denatured ubiquitin and comparisons with other denatured proteins. *J Biomol NMR*, 19(2):153–65.

Pike, A. C., Brew, K., and Acharya, K. R. (1996). Crystal structures of guinea-pig, goat and bovine alpha-lactalbumin highlight the enhanced conformational flexibility of regions that are significant for its action in lactose synthase. *Structure*, 4(6):691–703.

Piotto, M., Saudek, V., and Sklenar, V. (1992). Gradient-tailored excitation for single-quantum nmr spectroscopy of aqueous solutions. *J Biomol NMR*, 2(6):661–5.

Pometun, M. S., Peterson, R. W., Babu, C. R., and Wand, A. J. (2006). Cold denaturation of encapsulated ubiquitin. *J Am Chem Soc*, 128(33):10652–3.

Pons, J., Malliavin, T., and Delsuc, M. (1996). Gifa v. 4: A complete package for nmr data set processing. *J Biomol NMR*, 8:445–452.

Prestegard, J. H., Bougault, C. M., and Kishore, A. I. (2004). Residual dipolar couplings in structure determination of biomolecules. *Chem Rev*, 104(8):3519–40.

Prestegard, J. H., Valafar, H., Glushka, J., and Tian, F. (2001). Nuclear magnetic resonance in the era of structural genomics. *Biochemistry*, 40(30):8677–85.

Reckel, S., Lohr, F., and Dotsch, V. (2005). In-cell nmr spectroscopy. *Chembiochem*, 6(9):1601–6.

Redfield, A. G., Kunz, S. D., and Ralph, E. K. (1975). Dynamic range in fourier transform proton magnetic resonance. *J. Magn. Reson.*, 19(1):114–117.

Rehm, T., Huber, R., and Holak, T. A. (2002). Application of nmr in structural proteomics: screening for proteins amenable to structural analysis. *Structure*, 10(12):1613–8.

Rosen, M. K., Gardner, K. H., Willis, R. C., Parris, W. E., Pawson, T., and Kay, L. E. (1996). Selective methyl group protonation of perdeuterated proteins. *J. Mol. Biol.*, 263(5):627–36.

Ross, A., Salzmann, M., and Senn, H. (1997). Fast-hmqc using ernst angle pulses: An efficient tool for screening of ligand binding to target proteins. *J. Biomol. NMR*, 10:389.

Rossy, E., Champier, L., Bersch, B., Brutscher, B., Blackledge, M., and Coves, J. (2004). Biophysical characterization of the merp-like amino-terminal extension of the mercuric reductase from ralstonia metallidurans ch34. *J. Biol. Inorg. Chem.*, 9:49–58.

Rovnyak, D., Frueh, D. P., Sastry, M., Sun, Z.-Y. J., Stern, A. S., Hoch, J. C., and Wagner, G. (2004a). Accelerated acquisition of high resolution triple-resonance spectra using non-uniform sampling and maximum entropy reconstruction. *J Magn Reson*, 170(1):15–21.

Rovnyak, D., Hoch, J., Stern, A., and Wagner, G. (2004b). Resolution and sensitivity of high eld nuclear magnetic resonance spectroscopy. *J. Biomol. NMR*, 30:1–10.

Rubinstenn, G., Vuister, G. W., Mulder, F. A., Dux, P. E., Boelens, R., Hellingwerf, K. J., and Kaptein, R. (1998). Structural and dynamic changes of photoactive yellow protein during its photocycle in solution. *Nat Struct Biol*, 5(7):568–70.

Ruckert, M. and Otting, G. (2000). Alignment of biological macromolecules in novel nonionic liquid crystalline media for nmr experiments. *J. Am. Chem. Soc.*, 122:7793–7797.

Schanda, P. and Brutscher, B. (2005). Very fast two-dimensional nmr spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds. *J Am Chem Soc*, 127(22):8014–5.

Schanda, P. and Brutscher, B. (2006). Hadamard frequency-encoded sofast-hmqc for ultrafast two-dimensional protein nmr. *J Magn Reson*, 178(2):334–9.

Schanda, P., Kupce, E., and Brutscher, B. (2005). Sofast-hmqc experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds. *J Biomol NMR*, 33(4):199–211.

Schanda, P., Van Melckebeke, H., and Brutscher, B. (2006). Speeding up three-dimensional protein nmr experiments to a few minutes. *J Am Chem Soc*, 128(28):9042–3.

Schneider, D. M., Dellwo, M. J., and Wand, A. J. (1992). Fast internal main-chain dynamics of human ubiquitin. *Biochemistry*, 31(14):3645–52.

Selenko, P., Gregorovic, G., Sprangers, R., Stier, G., Rhani, Z., Kramer, A., and Sattler, M. (2003). Structural basis for the molecular recognition between human splicing factors u2af65 and sf1/mbbp. *Mol Cell*, 11(4):965–976.

Selenko, P., Serber, Z., Gadea, B., Ruderman, J., and Wagner, G. (2006). Quantitative nmr analysis of the protein g b1 domain in xenopus laevis egg extracts and intact oocytes. *Proc Natl Acad Sci U S A*, 103(32):11904–9.

Shapira, B., Karton, A., Aronzon, D., and Frydman, L. (2004a). Real-time 2d nmr identification of analytes undergoing continuous chromatographic separation. *J Am Chem Soc*, 126(4):1262–1265.

Shapira, B., Lupulescu, A., Shrot, Y., and Frydman, L. (2004b). Line shape considerations in ultrafast 2d nmr. *J Magn Reson*, 166(2):152–63.

Shapira, B., Morris, E., Muszkat, K., and Frydman, L. (2004c). Sub-second 2d nmr spectroscopy at sub-millimolar concentrations. *J Am Chem Soc*, 126(38):11756–11757.

Shrot, Y. and Frydman, L. (2003). Single-scan nmr spectroscopy at arbitrary dimensions. *J Am Chem Soc*, 125(37):11385–96.

Shrot, Y., Shapira, B., and Frydman, L. (2004). Ultrafast 2d nmr spectroscopy using a continuous spatial encoding of the spin interactions. *J Magn Reson*, 171(1):163–70.

Sibille, N., Blackledge, M., Brutscher, B., Coves, J., and Bersch, B. (2005). Solution structure of the sulfite reductase flavodoxin-like domain from escherichia coli. *Biochemistry*, 44(25):9086–95.

Simorre, J., Brutscher, B., Caffrey, M., and Marion, D. (1994). Assignment of nmr spectra of proteins using triple-resonance two-dimensional experiments. *J. Biomol. NMR*, 4(3):325–333.

Simorre, J. and Marion, D. (1991). A method aimed at obtaining a complete set of cross peaks in single-scan high-resolution homonuclear 3d nmr. *J Magn Reson*, 94(2):426–432.

Sivaraman, T., Arrington, C. B., and Robertson, A. D. (2001). Kinetics of unfolding and folding from amide hydrogen exchange in native ubiquitin. *Nat Struct Biol*, 8(4):331–3.

Sklenar, V., Pertersen, R., Rejante, M., and Feigon, J. (1993). Two- and three-dimensional hcn experiments for correlating base and sugar resonances in 15n,13c-labeled rna oligonucleotides. *J. Biomol. NMR*, 3(6):721–727.

Smith, M., Hu, H., and Shaka, A. (2001). Improved broadband inversion performance for nmr in liquids. *J. Magn. Reson.*, 151:269–283.

Song, X.-J., Flynn, P. F., Sharp, K. A., and Wand, A. J. (2007). Temperature dependence of fast dynamics in proteins. *Biophys J*, 92(6):L43–5.

States, D. J., Haberkorn, R. A., and Ruben, D. J. (1982). A two-dimensional nuclear overhauser experiment with pure absorption phase in 4 quadrants. *J. Magn. Reson.*, 48:286–292.

Staunton, D., Owen, J., and Campbell, I. D. (2003). Nmr and structural genomics. *Acc Chem Res*, 36(3):207–14.

Stonehouse, J., Shaw, G. L., Keeler, J., and Laue, E. D. (1994). Minimizing sensitivity losses in gradient-selected n-15-h-1 hsqc spectra of proteins. *J. Magn. Reson. A*, 107(2):178–184.

Szyperski, T., Wider, G., Bushweller, J., and Wuthrich, K. (1993a). 3d 13c-15n-heteronuclear two-spin coherence spectroscopy for polypeptide backbone assignments in 13c-15n-double-labeled proteins. *J. Biomol. NMR*, 3:127–132.

Szyperski, T., Wider, G., Bushweller, J., and Wüthrich, K. (1993b). Reduced dimensionality in triple-resonance nmr experiments. *J. Am. Chem. Soc.*, 115:9307–9308.

Tjandra, N. and Bax, A. (1997). Direct measurement of distances and angles in biomolecules by nmr in a dilute liquid crystalline medium. *Science*, 278(5340):1111–1114.

Tjandra, N., Feller, S., Pastor, R., and Bax, A. (1995). Rotational diffusion anisotropy of human ubiquitin from 15n nmr relaxation. *J. Am. Chem. Soc.*, 117:12562–12566.

Tossavainen, H. and Permi, P. (2004). Optimized pathway selection in intraresidual triple-resonance experiments. *J Magn Reson*, 170(2):244–51.

Tricot, C., van Aelst, S., Wattiez, R., Mergeay, M., Stalon, V., and Wouters, J. (2005). Overexpression, purification, crystallization and crystallographic analysis of copk of cupriavidus metallidurans. *Acta Crystallograph Sect F Struct Biol Cryst Commun*, 61(Pt 9):825–7.

Troullier, A., Reinstadler, D., Dupont, Y., Naumann, D., and Forge, V. (2000). Transient non-native secondary structures during the refolding of alpha-lactalbumin detected by infrared spectroscopy. *Nat Struct Biol*, 7(1):78–86.

Tugarinov, V., Choy, W.-Y., Orekhov, V. Y., and Kay, L. E. (2005). Solution nmr-derived global fold of a monomeric 82-kda enzyme. *Proc Natl Acad Sci U S A*, 102(3):622–7.

Tugarinov, V., Hwang, P. M., Ollerenshaw, J. E., and Kay, L. E. (2003). Cross-correlated relaxation enhanced 1h[bond]13c nmr spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *J Am Chem Soc*, 125(34):10420–8.

Ulmer, T. S. and Bax, A. (2005). Comparison of structure and dynamics of micelle-bound human alpha-synuclein and parkinson disease variants. *J Biol Chem*, 280(52):43179–87.

Urbauer, J., McEvoy, R., and Bieber, R. (1996). Internal dynamics of human ubiquitin revealed by 13c-relaxation studies of randomly fractionally labeled protein. *Biochemistry*, 35(19):6116 –6125.

Uversky, V. N. (2002). Natively unfolded proteins: a point where biology waits for physics. *Protein Sci*, 11(4):739–56.

Van Melckebeke, H., Simorre, J.-P., and Brutscher, B. (2004). Amino acid-type edited nmr experiments for methyl-methyl distance measurement in 13c-labeled proteins. *J Am Chem Soc*, 126(31):9584–91.

Van Nuland, N., Forge, V., Balbach, J., and Dobson, C. (1998). Real-time nmr studies of protein folding. *Acc. Chem. Res.*, 31(11):773–780.

Vanbelle, C., Brutscher, B., Blackledge, M., Muhle-Goll, C., Remy, M.-H., Masson, J.-M., and Marion, D. (2003). Nmr study of the interaction between zn(ii) ligated bleomycin and streptoalloteichus hindustanus bleomycin resistance protein. *Biochemistry*, 42(3):651–63.

Venters, R., Coggins, B., Kojetin, D., Cavanagh, J., and Zhou, P. (2005). (4,2)d projection–reconstruction experiments for protein backbone assignment: application to human carbonic anhydrase ii and calbindin d(28k). *J Am Chem Soc*, 127:8785–8795.

Vijay-Kumar, S., Bugg, C. E., and Cook, W. J. (1987). Structure of ubiquitin refined at 1.8 a resolution. *J Mol Biol*, 194(3):531–44.

Wagner, G. and Wuthrich, K. (1982). Amide protein exchange and surface conformation of the basic pancreatic trypsin inhibitor in solution. studies with two-dimensional nuclear magnetic resonance. *J Mol Biol*, 160(2):343–61.

Wand, A., Ehrhardt, M., and Flynn, P. (1998). High-resolution nmr of encapsulated proteins dissolved in low-viscosity fluids. *Proc. Natl. Acad. Sci. U.S.A.*, 95:15299–15302.

Wang, T., Cai, S., and Zuiderweg, E. R. P. (2003). Temperature dependence of anisotropic protein backbone dynamics. *J Am Chem Soc*, 125(28):8639–43.

Weigelt, J. (1998). Single scan, sensitivity- and gradient-enhanced trosy for multidimensional nmr experiments. *J. Am. Chem. Soc.*, 120:10778–10779.

Wenter, P., Furtig, B., Hainard, A., Schwalbe, H., and Pitsch, S. (2005). Kinetics of photoinduced rna refolding by real-time nmr spectroscopy. *Angew Chem Int Ed Engl*, 44(17):2600–3.

Wider, G. (1998). Technical aspects of nmr spectroscopy with biological macromolecules and studies of hydration in solution. *Prog. NMR Spectrosc.*, 32:193–275.

Wijesinha-Bettoni, R., Dobson, C. M., and Redfield, C. (2001). Comparison of the denaturant-induced unfolding of the bovine and human alpha-lactalbumin molten globules. *J Mol Biol*, 312(1):261–73.

Wilkinson, K. D. and Mayer, A. N. (1986). Alcohol-induced conformational changes of ubiquitin. *Arch Biochem Biophys*, 250(2):390–9.

Wimperis, S. and Bodenhausen, G. (1986). Heteronuclear coherence transfer over a range of coupling constants. *J Magn Reson*, 69:264–282.

Woessner, D. E. (1962). Spin relaxation processes in a two-proton system undergoing anisotropic reorientation. *J. Chem. Phys.*, 36:1–4.

Wolynes, P. G., Onuchic, J. N., and Thirumalai, D. (1995). Navigating the folding routes. *Science*, 267(5204):1619–20.

Wu, Z. and Bax, A. (2002). Measurement of long-range 1h-1h dipolar couplings in weakly aligned proteins. *J. Am. Chem. Soc.*, 124(33):9672–3.

Wuthrich, K., Otting, G., and Liepinsh, E. (1992). Protein hydration in aqueous solutions. *Faraday Disc.*, 93:35–45.

Zeeb, M. and Balbach, J. (2004). Protein folding studied by real-time nmr spectroscopy. *Methods*, 34(1):65–74.

Résumé

La RMN multidimensionnelle (RMN-nD) est la méthode de choix pour l'étude de la structure et de la dynamique des protéines en solution avec une résolution atomique. Une limitation de la RMN-nD est la longue durée de l'acquisition: le temps d'acquisition du jeu de données nécessaire pour une étude structurale est souvent de l'ordre de plusieurs semaines. De plus, des processus cinétiques, qui se passent à l'échelle de la seconde, ne sont pas accessibles aux études en temps réel par RMN-nD en utilisant les méthodes standards. Ce travail présente des développements méthodologiques qui visent à accélérer la RMN-nD en optimisant la relaxation longitudinale des protons amides. Les méthodes proposées permettent d'acquérir des spectres de corrélation 2D ^{1}H-^{15}N (3D ^{1}H-^{15}N-^{13}C) en quelques secondes (quelques minutes). En plus, en combinaison avec des méthodes existantes (encodage spatial, encodage Hadamard), le temps d'acquisition pour des spectres 2D peut être réduit à une seconde. Des applications à l'étude de phenomène cinétiques des protéines sont montrées.

Cette thèse présente aussi une nouvelle expérience RMN qui permet d'évaluer rapidement la qualité d'un échantillon de protéine, et une nouvelle méthode pour mesurer des couplages dipolaires résiduels entre protons amides avec une meilleure sensibilité que les méthodes existantes.

Abstract

Multidimensional (nD) NMR is the method of choice for atom-resolved studies of protein structure and dynamics in solution. Among its current limitations are the long acquisition times required, translating to experimental times of several days or weeks for the set of experiments required for structural studies of proteins. Furthermore, real-time studies of kinetic processes occurring on a seconds time scale are inaccessible to standard nD NMR. This thesis is concerned with the development of fast nD NMR techniques based on longitudinal relaxation optimization. It is shown that 2D ^{1}H-^{15}N (3D ^{1}H-^{15}N-^{13}C) correlation spectra can be obtained in only a few seconds (few minutes) of acquisition time for samples at millimolar concentration. In addition, the longitudinal relaxation optimized methods, when combined with alternative data sampling such as spatial or Hadamard encoding, can yield site-resolved 2D ^{1}H-^{15}N correlation spectra in acquisition times down to one second. Applications of fast 2D methods to the study of protein folding and unfolding are shown.

This thesis also presents a longitudinal relaxation optimized method for the sensitivity-enhanced measurement of residual dipolar couplings between amide protons, as well as a fast and simple experiment for the characterization of protein samples, which can be very useful in the context of screening of sample conditions.

Mots Clés: RMN multidimensionnelle, relaxation longitudinale, Effet Overhauser Nucléaire, Couplage Résiduel Dipolaire, protéine, repliement de protéines, échange H/D des protons amides

---

Laboratoire de thèse: Institut de Biologie Structurale Jean-Pierre Ebel, UMR 5075, CEA.; CNRS.; UJF. Laboratoire de Résonance Magnétique Nucléaire. 41, rue Jules Horowitz, 38027 Grenoble Cedex